

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ  
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ  
СІКОРСЬКОГО»  
ІНСТИТУТ ПРИКЛАДНОГО СИСТЕМНОГО АНАЛІЗУ  
КАФЕДРА МАТЕМАТИЧНИХ МЕТОДІВ СИСТЕМНОГО АНАЛІЗУ

На правах рукопису  
УДК 004.896

До захисту допущено  
В. о. завідувача кафедри ММСА

О.Л.Тимошук

«\_\_\_» \_\_\_\_\_ 2020 р.

**Магістерська дисертація**

на здобуття ступеня магістра за спеціальністю 124 Системний аналіз  
на тему: «Система підтримки прийняття рішень для оцінювання ризику  
акцій фінансового ринку»

Виконала:

студентка II курсу, групи КА-91мп

Пущик Оксана Романівна

\_\_\_\_\_

Керівник: професор кафедри ММСА

д.т.н, проф. Бідюк П. І.

\_\_\_\_\_

Рецензент:

Професор кафедри Інформаційної безпеки

КПІ ім. Ігоря Сікорського, д.т.н, проф. Архипов О.Є.

\_\_\_\_\_

Засвідчую, що у цій магістерській дисертації  
немає запозичень з праць інших авторів  
без відповідних посилань

Студент \_\_\_\_\_

Київ  
2020

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ  
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ  
СІКОРСЬКОГО»  
ІНСТИТУТ ПРИКЛАДНОГО СИСТЕМНОГО АНАЛІЗУ  
КАФЕДРА МАТЕМАТИЧНИХ МЕТОДІВ СИСТЕМНОГО АНАЛІЗУ

Рівень вищої освіти — другий (магістерський)  
Спеціальність — 124 «Системний аналіз»

ЗАТВЕРДЖУЮ

В. о. завідувача кафедри ММСА

О. Л. Тимошук

«\_\_\_» \_\_\_\_\_ 2020 р.

### ЗАВДАННЯ

на магістерську дисертацію студентці Пущик Оксані Романівні

**1. Тема дисертації:** «Система підтримки прийняття рішень для оцінювання ризику акцій фінансового ринку», науковий керівник дисертації Бідюк Петро Іванович, д.т.н., професор, затверджені наказом по університету від «02» листопада 2020 р. № 3182-с.

**2. Термін подання студентом дисертації:** 14 грудня 2020 р.

**3. Об'єкт дослідження:** Нестационарні фінансово-економічні процеси зі змінною у часі волатильністю.

**4. Предмет дослідження:** математичні моделі і методи опису гетероскедастичних процесів, методи прогнозування часових рядів, оцінювання та аналізу якості побудованих моделей та прогнозів, моделі та методи оцінювання ринкових ризиків, а також методи перевірки якості оцінок ризику.

**5. Перелік завдань, які потрібно розробити:**

- 1) дослідити сучасний стан та особливості застосування машинного навчання у вирішенні проблеми оцінювання ризиків акцій;
- 2) розробити математичну модель нейронної мережі для прогнозування волатильності фінансових процесів;
- 3) розібрати і підібрати моделі для прогнозування гетероскедастичних часових рядів;
- 4) на основі моделей створити програмний продукт для визначення ризику акцій фінансового ринку;
- 5) знайти дані для застосування в програмі;
- 6) розробити стартап-проект виведення на ринок результатів дослідження;

7) розробити концептуальні висновки за результатами наукового дослідження.

**6. Орієнтовний перелік графічного (ілюстративного) матеріалу:**

- 1). Огляд існуючих нейронних мереж для прогнозу волатильності (рис.2.1 – рис.2.3);
- 2). Робота створеного програмного продукту (рис.3.1 - рис.3.3);
- 3). Таблиці у розділі стартап-проекту.

**7. Орієнтовний перелік публікацій:**

- (1) Пущик О. Р. Оцінювання волатильності акцій фондового ринку. Проміжні результати. 2020. DOI: <http://dx.doi.org/10.13140/RG.2.2.26668.08325>

**8. Дата видачі завдання:** 1 вересня 2019 року.

**Календарний план**

№ з/п	Назва етапів виконання магістерської дисертації	Термін виконання етапів магістерської дисертації
1.	Концептуальний вступ дисертації. Формулювання об'єкта, предмета, цілі, завдань, новизни, практичної значущості результатів.	1.09.2020 – 5.09.2020
2.	Огляд літературно-інформаційних джерел.	5.09.2020 – 10.09.2020
3.	Перший розділ. Понятійно-категоріальний апарат. Характеристика об'єкта.	10.09.2020 – 25.09.2020
4.	Другий розділ. Детальний розгляд систем для трейдингу і процесів на біржі.	25.09.2020 – 30.09.2020
5.	Третій розділ. Імплементация отриманих результатів у програмний продукт. Тестування програми.	30.09.2020 – 25.10.2020
6.	П'ятий розділ. Стартап-проект.	25.10.2020 – 20.11.2020
7.	Концептуальні висновки. Перспективи розвитку отриманих рішень.	20.11.2020 – 9.12.2020

Студент

Пущик О. Р.

Науковий керівник дисертації

Бідюк П. І.

## РЕФЕРАТ

Магістерська дисертація: 111 с., 7 рис., 27 табл., 1 додаток, 27 джерел.

РИНКОВИЙ РИЗИК, ВОЛАТИЛЬНІСТЬ, ПРОГНОЗУВАННЯ, УМОВНА АВТОРЕГРЕСІЙНА ГЕТЕРОСКЕДАСТИЧНІСТЬ, РЕКУРЕНТНІ НЕЙРОННІ МЕРЕЖІ, РИЗИК АКЦІЙ, VALUE-AT-RISK.

Об'єкт дослідження – нестационарні фінансово-економічні процеси зі змінною у часі волатильністю.

Предмет дослідження – математичні моделі і методи опису гетероскедастичних процесів, методи прогнозування часових рядів, оцінювання та аналізу якості побудованих моделей та прогнозів, моделі та методи оцінювання ринкових ризиків, а також методи перевірки якості оцінок ризику.

Методи дослідження – теорія моделювання і прогнозування, регресійний аналіз, статистичні методи.

Метою роботи є побудова системи підтримки прийняття рішень, яка включає в себе адекватну модель гетероскедастичного процесу для прогнозування волатильності та оцінювання ризику акцій фінансового ринку за її допомогою.

В роботі проведено огляд основних підходів до оцінювання ринкових ризиків, розглянуто та проаналізовано метод оцінки Value-at-Risk. Також проведений огляд моделей та їх особливостей для опису динаміки волатильності та її прогнозування. Було проаналізовано результати моделювання та оцінювання за-для обґрунтованого вибору найкращої моделі для оцінки ринкових ризиків. Моделювання процесів на базі авторегресійних умовно гетероскедастичних моделей та на базі рекурентних нейронних мереж для оцінювання ризикової вартості за їх допомогою реалізовано на мові програмування Python.

## ABSTRACT

Master thesis: 111 p., 7 fig., 27 tab., 1 application, 27 sources.

MARKET RISK, VOLATILITY, FORECASTING, CONDITIONAL AUTOREGRESSIONAL HETEROSCEDASTICITY, RECURRENT NEURAL NETWORKS, SHARE RISK, VALUE-AT-RISK.

Object of study - non-stationary financial-economic processes with timevarying volatility.

Subject of investigation - mathematical models and methods for describing heteroscedastic processes, time series forecasting methods, evaluating and analyzing the quality of the models and forecasts, models and methods for estimation of market risk, and methods for backtesting of risk estimates.

Methods - Theory of modeling and forecasting, regression analysis, statistical methods.

The aim is to build a decision support system that includes an adequate model of the heteroskedastic process for forecasting volatility and assessing the risk of financial market shares with its help.

In this paper, it is reviewed of the main approaches to market risk estimation, reviewed and analyzed the method for estimating Value-at—Risk and applied innovative methods for verifying the quality of these estimates. Also reviewed models and their features to describe the dynamics of volatility and its forecasting. Results of modeling, forecasting and evaluation were analyzed for selecting the best model for market risks estimation.

Modeling and forecasting of financial and economic processes on the basis of autoregressive conditionally heteroscedastic models and recurrent neural networks estimating the risk with their help are implemented in the programming language Python.

## ЗМІСТ

СПИСОК СКОРОЧЕНЬ.....	8
ВСТУП .....	9
РОЗДІЛ 1. ТЕОРЕТИЧНІ ОСНОВИ ПРО РИНКОВІ РИЗИКИ І СИСТЕМИ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ.....	12
1.1 Прийняття рішень.....	12
1.2 Система підтримки прийняття рішень .....	14
1.3 Поняття фондового ринку і акцій .....	17
1.4 Ринкові ризики .....	19
1.5 Ризики акцій фінансового ринку.....	21
1.6 Методи керування ринковими ризиками .....	22
1.7 Підходи до вимірювання ризику.....	23
1.8 Міри ризику на основі розподілу втрат.....	25
1.9 Зашумованість даних.....	28
1.10 Методи очищення даних від шумів .....	29
1.11 Висновки до розділу .....	30
РОЗДІЛ 2. ОПИС МОДЕЛЕЙ ДЛЯ ОЦІНКИ РИЗИКУ АКЦІЙ.....	32
2.1 Фільтр Калмана і його модифікації.....	32
2.2 Моделі для оцінювання часових рядів.....	37
2.3 Нейронні мережі.....	42
2.4 Перевірка наявності гетероскедастичності .....	47
2.5 Методи перевірки моделі на адекватність.....	49
2.6 Критерії якості оцінок прогнозів.....	51
2.7 Висновки до розділу .....	54
РОЗДІЛ 3. АНАЛІЗ ОБЧИСЛЮВАЛЬНИХ ЕКСПЕРИМЕНТІВ ТА ОГЛЯД ПРОГРАМНОГО ПРОДУКТУ .....	55
3.1 Вхідні дані .....	55
3.2 Результати оцінювання .....	56
3.3 Функціональна схема алгоритму .....	60
3.4 Висновки до розділу .....	61

РОЗДІЛ 4. РОЗРОБКА СТАРТАП ПРОЕКТУ .....	63
4.1 Постановка задачі стартап-проекту.....	63
4.2 Карта проекту .....	64
4.3 Команда стартап- проекту.....	66
4.4 Розробка бізнес-моделі проекту.....	68
4.5 Аналіз ринкових можливостей запуску стартап-проекту .....	72
4.6 Розроблення ринкової стратегії проекту.....	82
4.7 Розробка маркетингової програми стартап-проекту.....	87
4.8 Висновки до розділу .....	91
ВИСНОВОК.....	93
ПЕРЕЛІК ПОСИЛАНЬ.....	95
ДОДАТОК А. ЛІСТИНГ ПРОГРАМИ .....	98

## СПИСОК СКОРОЧЕНЬ

AP – авторегресія.

APКС – авторегресія з ковзним середнім.

КС – ковзне середнє.

СППР – система підтримки прийняття рішень.

ARCH – авторегресійна умовна гетероскедастичність.

GARCH – узагальнена авторегресійна умовна гетероскедастичність.

NN – нейронна мережа.

RNN – рекурентна нейронна мережа.

LSTM – мережа довгої короткотривалої пам'яті.

VaR – Value-at-Risk – вартість під ризиком.



## ВСТУП

Магістерська робота присвячена оцінці фондових ризиків і методикам, що дозволяють це зробити.

На даний момент йде постійне зростання фінансових і, особливо, фондових ринків. Даний зріст привертає увагу великої кількості людей, які хочуть отримати прибуток на цих платформах. Відповідно це привело до різкого росту кількості методологій для роботи з подібними даними. Оскільки традиційні методи роботи з часовими рядами далеко не завжди дають хороші результати. В світі постійно вивчаються фінансові часові ряди, і за рахунок цього було виявлено уже певні особливості таких даних. В основному фінансові дані наділені властивостями довгої пам'яті, відсутністю автокореляції, дуже часто присутня гетероскедастичність даних. Це все призвело до появи нових методів роботи з фінансовими ринками. Серед таких моделей є модель умовної гетероскедастичності (ARCH) і узагальнена модель умовної гетероскедастичності (GARCH), а також їх варіації.

Крім того розвиваються і інші методики роботи з даними залежними від часу. Наприклад, рекурентні нейронні мережі. Весь цей набір методологій можна використовувати і для даних фінансового ринку.

Ще один момент, який зараз хвилює всіх підприємців – це ризики і наслідки від них, які вони можуть понести. Зараз в компаніях зазвичай існують цілі відділи, що займаються подібними питаннями, проте є ризики, які досить складно передбачити. Одними з таких ризиків є випадкові зміни цін на акції, на курс валют, на процентні ставки, тощо. На даний момент існує уже багато методик оцінки даних ризиків: Value-at-risk, Expected Shortfall, Capital-at-Risk, Maximum Loss, тощо. Найвідомішим і найвикористованішим серед них є показник VaR, оскільки він досить легко

інтерпретується (по факту є оцінкою можливих максимальних втрат при певній ймовірності), є легким у використанні. І досить легко обчислюється.

Одною із найцікавіших областей застосування даних методологій є фондові ринки. Існує багато бажаючих отримати прибуток нічого не роблячи, проте насправді робота трейдерів є досить складною в плані оцінки, що краще купити, що продати, а що і притримати поки ще не зросте ціна на акції. Одною з найбільших проблем даної області є нестабільність ринку. Ціни акцій постійно змінюються, і причина може бути різною: економічна, політична, соціальна і т.д. Проте вміння вчасно зреагувати на ризики і прийняти рішення про купівлю-продаж може привести до значного власного збагачення. Тому у даній роботі пропонується застосування VaR для оцінки ризиків акцій фінансового ринку.

Об'єкт дослідження – нестационарні фінансово-економічні процеси фондового ринку зі змінною у часі волатильністю.

Предмет дослідження – моделі та методи для оцінки і прогнозу нестационарних процесів.

Мета роботи – розробити систему підтримки прийняття рішень для оцінки ризиків акцій фондового ринку.

Актуальність роботи пов'язана з потребою в оцінюванні ринкових ризиків для подальшого прийняття адекватного, обґрунтованого рішення. Дана робота знаходить практичне застосування, оскільки дозволяє проведенню обдуманих операції на біржі.

У першому розділі розглянуто основні теоретичні аспекти роботи з фінансовими процесами і зокрема оцінкою їх ризиків. Приводиться опис методів оцінювання ринкових ризиків і керування даними ризиками.

У другому розділі розповідається про основні методи визначення показника VaR, а також моделей для прогнозування волатильності часового ряду. Розглянуто два основних напрямки прогнозування часових рядів: моделі умовної гетероскедастичності та рекурентні нейронні мережі. Також приводиться теоретична база для методів видалення шумів з даних.

В третьому розділі проводиться розгляд реалізації результатів отриманих в ході експериментів. Представлено отримані результати і ефективність методів для конкретних даних. Також розглянуто і представлено схему розробленої СППР.

Четвертий розділ присвячено розробці стартап-проекту на базі отриманих результатів.

У висновках представлено основні думки і підсумки по даній роботі.

## РОЗДІЛ 1. ТЕОРЕТИЧНІ ОСНОВИ ПРО РИНКОВІ РИЗИКИ І СИСТЕМИ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ

### 1.1 Прийняття рішень

Прийняття рішень – це вибір варіанта, який підходить найкраще з усіх можливих варіантів для конкретної ситуації чи задачі.

Теорію прийняття рішень можна застосовувати до об'єктів різноманітної природи і при різноманітних умовах їх існування. При цьому процеси прийняття рішень в різних сферах людської діяльності мають дуже багато спільного.

Формальні методи прийняття рішень можуть виявитись корисними в таких ситуаціях:

1. Існує певна проблема або проблемна ситуація, що потребує свого власного вирішення. Часто результат, який хочуть отримати, збігається з цілями, які повинні бути досягнені при розв'язку проблеми.

2. Існує кілька варіантів розв'язку проблеми, способів досягнення цілей, серед яких вибирається оптимальний варіант. Ці варіанти називаються альтернативами. Якщо існує одна можливість і вибору немає, то немає і самої задачі прийняття рішень.

3. Присутні фактори, які надають певні обмеження на можливі шляхи розв'язку проблеми. Ці фактори визначаються контекстом проблеми, залежать від конкретної ситуації і можуть бути різної природи: фізичної, економічної, соціальної, технічної, тощо.

4. Існує група людей, які зацікавлені в розв'язку даної задачі, вони мають повноваження для вибору певного рішення і несуть відповідальність за виконання даного рішення.

Сам процес прийняття рішень має певну типову схему. Тут існує кілька стадій і разом вони є ітеративною процедурою.

Необхідність прийняття рішень виникає при появі проблемної ситуації. В цьому випадку спершу необхідно виявити саму проблему, тобто описати її, а потім визначити бажаний результат при її рішенні. Оцінити всі можливі обмеження, які присутні в даній проблемі.

На наступній стадії відбувається постановка задачі прийняття рішень. Для цього необхідно визначити сукупність всіх можливих варіантів розв'язку даної задачі. Тобто визначити всі можливі альтернативи. В залежності від контексту задачі і її особистих характеристик кількість альтернатив може бути різною – від пари альтернатив до кількох тисяч, тощо. Теоретично кількість альтернатив може бути безкінечною. Для визначення альтернатив зазвичай необхідно зібрати інформацію по даній проблемі і проаналізувати її.

В кінці кінців необхідно отримати постановку задачі прийняття рішень.

Наступний етап – це пошук рішення проблеми. Проте не завжди можна знайти найкраще рішення. В цьому випадку необхідно переглянути заново постановку задачі і ввести корективи до неї.

Фінальний етап – це втілення в життя рішення проблеми, що приведе до її зникнення. Проте завжди треба оцінювати кінцевий результат, адже вирішивши одну проблему, можуть з'явитись нові. При чому їх поява може бути спричинена саме нашим рішенням попередньої проблеми.

Іноколи трапляються ситуації, коли формально описати задачу неможливо, тоді в гру вступають спеціалісти різноманітних галузей знань. Це називається проблемою прийняття рішень.

З цього слідує, що сама постановка задачі сама стає проблемою, і для неї також необхідно придумувати різні методи для вирішення. В фіналі ми повинні отримати вираз, який би описував нашу проблему формально, при цьому зв'язуючи ціль і методи її досягнення.

## 1.2 Система підтримки прийняття рішень

Задача прийняття рішень зазвичай дуже складні і об'ємні. Тому для роботи з ними існує інструмент, що допомагає знайти рішення проблеми. Цим інструментом є система підтримки прийняття рішень. Це інформаційна система, яка використовує обладнання, програмне забезпечення, дані, базу моделей з метою підтримки всіх стадій прийняття рішень у процесі моделювання задачі.

Сучасні СППР з'явилися внаслідок злиття інформаційних систем і систем керування базами даних.

Відповідно до Турбана [1, 2] ідеальній системі підтримки прийняття рішень притаманні такі характеристики:

- оперує зі слабо структурованими рішеннями;
- призначена для користувачів різного рівня;
- може бути застосована для групового та індивідуального використання;
- підтримує послідовні та взаємозалежні рішення;
- підтримує інтелектуальну частину, проектування та вибір рішення;
- підтримує різні методи рішення;
- гнучка і адаптується до змін;
- проста у використанні;
- підвищує ефективність процесу прийняття рішення;
- дозволяє людині керувати процесом за допомогою комп'ютера;
- легко адаптується до вимог, що можуть змінюватись;
- легко будується;
- підтримує моделювання;

— дозволяє використовувати знання отримані в процесі і в фіналі роботи системи.

Проте далеко не завжди можливо створити ідеальну СППР, оскільки задачі керування можуть бути:

— дуже складними (мати величезну кількість факторів, що на неї впливають);

— мати зростаючу динаміку.

В таких ситуація можна використовувати інтелектуальні системи підтримки прийняття рішень (ІСППР). Загальна структура таких систем представлена на Рисунку 1.1.



Рисунок 1.1 – Загальна структура ІСППР

В даній схемі інтерфейс дозволяє користувачу зручно спілкуватись із системою і отримувати рекомендації в красивій, зрозумілій і зручній формі.

Підсистема аналізу проблем забезпечує класифікацію проблем та аналіз можливостей використання різних програмних продуктів для рішення проблеми.

Підсистема прийняття рішень забезпечує визначення альтернатив або рекомендацій для прийняття рішення.

Бази даних, знань і моделей призначені для збереження основних даних, що дозволяють працювати підсистемі прийняття рішень.

Системи підтримки прийняття рішень бувають різними. Причому різниця може бути на різних рівнях.

На рівні користувача СППР бувають пасивні (допомагаю процесу прийняття рішень, не вносить пропозицію для її прийняття), активні (здійснює пропозицію, яке рішення слід обрати), кооперативна (дає можливість змінювати, поповнювати та покращувати рішення, тобто кінцеве рішення – це узгоджене рішення між користувачем і системою).

На концептуальному рівні розділяють СППР на такі класи:

- орієнтовані на дані;
- орієнтовані на моделі;
- орієнтовані на знання;
- орієнтовані на документи;
- орієнтовані на комунікації;
- групові СППР;
- інтерорганізовані та інтраорганізовані;
- спеціалізовані СППР або СППР загального призначення;
- СППР на базі Web.

Крім того ще можна поділяти за ступенем обробки даних:

1. Інформаційно-пошукові. Системи шукають необхідні дані.
2. Оперативно-аналітичні. Такі системи допомагають групувати і узагальнювати дані.



3. Інтелектуальні. Дані СППР шукають логічні закономірності в даних, будують моделі, що описують процеси і прогнозують подальший розвиток процесу з певною ймовірністю.

В залежності від специфіки задач і призначення СППР кожен розробник вибирає оптимальний варіант системи, яка найкраще дозволила б вирішити задачу.

### 1.3 Поняття фондового ринку і акцій

Фондовий ринок – це фінансово-економічний інструмент світового масштабу, основна функція якого заключається в регулюванні оборту цінних паперів шляхом купівлі, продажу або застави. На фондових ринках торгують акціями, облігаціями і векселями.

Багато компанії, що тільки починають свій життєвий цикл не мають достатньо капіталу, а розвиток прогресивних ідей і технологій вимагають величезних сум грошей. В таких випадках перед стартом бізнес-проекту проводиться випуск цінних паперів – акцій, які купуються інвесторами і надії на ріст вартості і щорічні дивіденди.

Дуже часто на певному етапі розвитку комерційне підприємство випускає первинні акції, і пропонує до продажу додаткові обсяги акційних активів.

Фондовий ринок дає хороші можливості для заробітку грошей і розвитку нових компаній, технологій, ідей. Наприклад:

1. Отримання стабільного доходу від щорічних дивідендів після купівлі акцій.
2. Еволюція перспективних комерційних проектів з привабленням коштів інвесторів.

3. Отримання прибутку за схемою купівля-продаж цінних паперів за рахунок зміни котирувань (трейдинг).

4. Надання власних капіталів в довірче керування успішним брокерам і трейдерам з отриманням відповідної винагороди.

5. Додаткова емісія цінних паперів, що дозволяє розвинутим компаніям і проектам вийти на принципово новий рівень розвитку і прибутку.

Сам фондовий ринок ділиться на три великих сегмента для роботи з ним:

Первинний ринок. Продаж акцій молодими компаніями.

Вторинний ринок. Активна купівля-продаж акцій.

Позабіржовий ринок. Це ринок для трейдерів і брокерів. Вони заключають між собою договори про купівлю-продаж.

На фондових ринках торгують цінними паперами. Вони бувають трьох видів: акції, облігації, векселя.

Акції являють собою найпопулярніший цінний папір, який підтверджує право власника на частину майна відповідного підприємства з отриманням дивідендів від прибутку компанії.

Вони бувають двох видів:

1. Непривілейовані акції. Такі акції надають власнику право участі в загальних зборах акціонерів і отримання дивідендів.

2. Привілейовані акції. Вони зазвичай розповсюджуються серед засновників компанії і їх близьких. Вони дають можливість отримувати дивіденди і претендувати на власність ліквідованих компаній в пріоритетному порядку.

На сьогоднішній день на фондовому ринку є такі основні стратегії для заробітку від купівлі-продажі цінних паперів навіть з невеликим стартовим капіталом.

1. Пошук недооцінених акцій з інвестиціями в перспективу росту вартості.

2. Проведення договорів з участю коефіцієнта ціна/прибуток, що являє собою суму в доларах, яку трейдер повинен витратити для отримання одного долара чистого прибутку.

3. Стратегія swing trading, суть таких операцій в знаходженні загального тренду ринкових операцій щодо росту чи падіння і оперативним заключенням договору на малих позиціях.

4. Використання контрактів CFD, що заключаються на різницю в ціні.

Проте не потрібно забувати, що кожна операція на фондових ринках несе в собі ризики втрат. І основною задачею гравців на біржі є зведення цих ризиків до мінімуму при максимально можливих прибутках. Тому вміння проаналізувати дані і вибрати оптимальну стратегію є важливим аспектом роботи на фондових ринках.

#### 1.4 Ринкові ризики

Ризик – це ймовірність виникнення збитків або додаткових витрат, або недотримання запланованих доходів, або невиконання своїх обов'язків внаслідок негативних внутрішніх та зовнішніх подій.

Ринковий ризик – наявний або потенційний ризик для надходжень і капіталу, який виникає через несприятливі коливання вартості цінних паперів, товарів і курсів іноземних валют за тими інструментами, які є в торговельному портфелі. Цей ризик впливає з маркетмейкерства, дилінгу, прийняття позицій з боргових та пайових цінних паперів, валют, товарів та похідних інструментів (деривативів) [3].

Ринковий ризик має макроекономічну природу, тобто джерелом ринкового ризику є макроекономічні показники фінансової системи: індекси, процентні ставки, тощо.

Ринкові ризики можуть характеризуватись:

- попит і пропозиція на актив є основною причиною змін в ринкових цінах;
- неоднорідністю, тобто ринкові ризики також залежать і від фінансового стану емітента: рентабельність, прибутковість, тощо;
- взаємозалежністю складових елементів;
- інтегрованістю майже в усі види операцій фінансового ринку.

Всього існує чотири стандартні форми ринкових ризиків:

#### 1. Фондовий ризик.

Даний ризик несуть компанії, тому що вартість її цінних паперів падає. Об'єктами оцінки фондового ризику вважаються: акції і облігації, депозитарні розписки, деривативи, в основі яких лежать цінні папери підприємства.

#### 2. Процентний ризик.

Такий ризик несуть фінансові установи при несприятливих змінах процентних ставок. Наприклад: ринкова вартість ринкова вартість облігації з постійним купоном падає при підвищенні ставок прибутковості, отже, компанія-емітент позбавляється частини прибутку.

#### 3. Валютний ризик.

Валютний ризик - це ризик несприятливої зміни курсів іноземних валют.

#### 4. Товарний ризик.

Ризик зміни цін товарів.

Товарний ризик - це ризик, що визваний невизначеністю щодо майбутніх ринкових величин і розміру майбутнього прибутку, визваного коливаннями цін на товари.

Ринкові відносини не існують без ризиків. Тому основна задача всіх фінансових закладів відслідковування ризиків і знаходження оптимального методу управління ними.

## 1.5 Ризики акцій фінансового ринку

Цінні папери є чудовим способом забезпечити собі непоганий пасивний дохід. Акціонери мають прекрасну можливість отримувати процент від діяльності великих компаній, приймати участь і зборах, що присвячені аналізу фінансової діяльності компанії, вибору стратегічно важливого напрямку розвитку.

Акції бувають звичайними і привілейованими. Звичайні акції купуються з метою накопичення коштів і їх захисту від інфляції. Акції можуть змінюватись в ціні внаслідок змін на фондовому ринку, їх можна продати чи обміняти. Зберігання даних цінних паперів не гарантує отримання прибутку: все залежить від успішності діяльності компанії.

Привілейовані папери - це гарантовані дивіденди, які виплачуються завжди, не залежачи від прибутку компанії. Звичайні акції дають можливість для виплат після загальних зборів акціонерів. Такий ризик передбачає прийняття обдуманого рішення, щоб в результаті отримати продуману інвестиційну діяльність.

При цьому покупка акцій пов'язана з певними ризиками неотримання того проценту прибутку, на який розраховує акціонер. Фінансові кризи, політичні потрясіння і тому подібне здатні привнести в діяльність компанії корективи.

Ризики акцій фінансового ринку відносяться до фондових ризиків. Ризик акцій - це ризик, який характеризується зміною ціни на акції певної компанії.

## 1.6 Методи керування ринковими ризиками

Сучасна ринкова економіка вимагає посиленого значення правильного аналізу і оцінювання ризику.

Керування ризиком передбачає максимально повне розуміння ступеню ризику, який загрожує. Методів керування ризиками існує надзвичайно багато.

В основному системи керування ризиком зводяться до наступних етапів:

1. Ідентифікація. Тобто необхідно виявити потенційний ризик.
2. Контроль ризику. Сюди входять певні стандарти кожної конкретної компанії для керування ризиків, які включають специфіку даного підприємства.
3. Моніторинг ризику. Кожна організація намагається регулярно відстежувати рівні ризиків щоб своєчасно втрутитись і зменшити його рівень.

Зазвичай для контролю над ризиками організовується окремий підрозділ, який займається саме питаннями ризику.

Суть ринкових ризиків в тому, що з часом вони призводять до нестабільності грошових потоків, що дуже впливає на фінансовий стан компанії. Щоб вберегтись від цього існують механізми, що дозволяють керувати цими ризиками, тим самим забезпечуючи певну стійкість грошових потоків, а отже, і компанії.

Основними способами керування ризиками є:

1. Уникнення можливих ризиків.  
Даний спосіб означає відмовитись від ненадійних сценаріїв, проектів, партнерів.
2. Хеджування.

В даному методі можна уникнути втрат при укладанні угоди, в цьому випадку ризик переходить до іншої особи. Проте дана методика потребує додаткових втрат.

### 3. Резервування.

При цьому формується власний капітал для компенсацій потенційних втрат. Такий метод не зменшує ймовірність настання ризику, але допомагає в випадку його настання відшкодувати матеріальні збитки.

### 4. Диверсифікація.

Суть даного методу в зниженні максимально можливих втрат для однієї події. Зазвичай цей метод використовується у вигляді покупки більше ніж одного виду активів, при чому прибутки їх повинні бути не пов'язані між собою або хоча б слабо зв'язані.

Для ефективного керування ризиками необхідно здійснювати керування на всьому життєвому циклі проекту. На основі повторних аналізів ризику можна скорегувати або змінити стратегію керування ризиками.

## 1.7 Підходи до вимірювання ризику

Одним з перших, хто запропонував підхід оцінювання ринкових ризиків був Мілтон Фрідмен. Він запропонував оцінювання ризику за допомогою теорії корисності. Фрідмен акцентував увагу на неможливості використання принципів максимізації при спадній корисності та наявності ризику. Оскільки в цьому випадку потребується компенсації за фактори ризику.

Він вважав, що економічну одиницю можна подати у вигляді певної функції, що дозволить дати чисельні значення альтернативам.

Існують п'ять основних підходів до оцінки ринкових ризиків:

1. Комплексний підхід. Суть в тому, що йде оцінка впливу ризику на роботу підприємства, а отже, і на його основні показники.

2. Систематичний підхід. Йде оцінка коефіцієнту чутливості ризику до результатів роботи ринку або економіки. Даний метод має важливе значення саме при оцінці ризику акцій, оскільки допомагає зрозуміти наскільки очікуваний прибуток компенсує ризик вкладень в акції певної компанії. На основі цих показників можна оцінити, які акції краще купувати, які приносять максимальний дохід при мінімальному ризику.

3. Статистичний підхід. Суть даного підходу в оцінці ймовірності виникнення втрат на базі статистичних даних. Дані методи хороші за рахунок того, що вони дають можливість проаналізувати різні варіанти розвитку подій, а також включати певні характеристики і чинники ризику. В цьому підході використовуються методи: оцінка ймовірності настання ризику, імітаційне моделювання, тощо.

4. Аналітичні методи. Суть у використанні математичних моделей, які дозволяють оцінити ймовірність втрат.

5. Методи експертних оцінок. Суть даного підходу в опитуванні групи експертів і подальша обробка даної інформації за допомогою математично-статистичних методів. Тобто даний підхід дозволяє включити в розгляд інтуїцію, знання та досвід учасників опитування.

Суть роботи фінансових ризиків в співвідношенні ризику та доходів. Оцінювання ризиків з допомогою будь-якого підходу зосереджена саме на завданні знаходження оптимального варіанту, коли дохід максимальний при мінімальному ризику.

В основному для оцінки ринкових ризиків використовують доходність активу, яку можна розрахувати за формулою:

$$r_i = \frac{P_i - P_{i-1}}{P_i}, \quad (1.1)$$



де  $r_i$  — доходність активу на поточний (і-тий) момент часу;

$P_i$  і  $P_{i-1}$  — вартість активу відповідно на поточний (і-ий) та попередній (і-1-ий) моменти часу.

У випадку довгострокових періодів краще використовувати геометричну доходність за формулою:

$$r_i = \ln \frac{P_i}{P_{i-1}} \quad (1.2)$$

де  $r_i$  — доходність активу на поточний (і-тий) момент часу;

$P_i$  і  $P_{i-1}$  — вартість активу відповідно на поточний (і-ий) та попередній (і-1-ий) моменти часу.

Надзвичайно важливими характеристиками ринкових ризиків є волатильність (мінливість показників) і чутливість факторів ризику.

Волатильність — це міра мінливості фінансового активу. Тобто показник, того наскільки змінюються котирування акцій протягом певного проміжку часу.

Загальний принцип волатильності — чим вища волатильність, тим вищі ризики. Проте чим вищі ризики, тим вища доходність.

Аналіз чутливості — це метод оцінки впливу основних параметрів фінансової моделі на результуючий показник.

## 1.8 Міри ризику на основі розподілу втрат

Одним з найбільш відомих в інвестиційному аналізі методів оцінки ризиків — це показник середньоквадратичного відхилення доходу. Даний показник — це статистична характеристика волатильності.

Даний індикатор характеризує розмір коливань доходності відносно очікуваної доходності.

Даний показник можна розраховувати за ціною акцій або ж за доходністю. При чому другий варіант найбільш поширений в практиці, оскільки показник доходності є адекватнішим для порівняння характеристикою.

Як правило доходність визначається за ретроспективою. При цьому вважається, що дані незалежні та мають однаковий розподіл. Очікувану доходність можна знайти за підрахунком простого середнього, тоді ризик – це оцінка дисперсії. Тоді волатильність визначається як квадратний корінь із оцінки нашого ризику, тобто із оцінки дисперсії доходності за такою формулою:

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (r_i - \underline{r})^2}, \quad (1.3)$$

де  $N$  – кількість спостережень,

$\underline{r}$  – середнє значення, очікувана доходність.

Даний метод оцінки волатильності має значний недолік: оцінка дисперсії не розрізняє позитивні і негативні відхилення від очікуваної доходності. Тому дану міру можна використовувати лише для симетричних розподілів (наприклад: нормального). Проте це рідко зустрічається на практиці.

Ще одною важливою мірою для оцінювання ризиків акцій є Value-at-Risk. Дана міра оцінює величину втрат з заданою ймовірністю.

Кінцева величина ризику формується рівнем схильності до ризику і оцінкою ризику в формі можливих втрат і в кінці зводить ці дві характеристики в один показник.

Даний показник характеризується трьома параметрами:

1. Часовий горизонт. Він залежить від ситуації, що розглядається. Зазвичай використовують часовий горизонт в один день. Проте по Базелю прийнято брати десять днів.
2. Довірчий інтервал. Це рівень допустимого ризику. По Базелю потрібно використовувати 99%, проте часто на практиці використовується 95%.
3. Базова валюта. Параметр, що вказує на валюту, в якій вимірюється показник.

По суті характеристика VaR – це величина збитків, що з ймовірністю рівною рівню довіри не буде вищою.

Даний показник можна пояснити як: «Існує ймовірність на  $X\%$ , що втрати на перевищать  $Y$  доларів протягом наступних  $N$  днів». В даному випадку величина  $Y$  і є показником VaR.

В даній мірі є великий ряд переваг:

1. Легкість в інтерпретації.
2. Зручність в представленні навіть з іншими показниками.
3. Можливість використовувати в якості способу регулювання обсягів операцій відповідно до величини власних коштів.
4. Зручно і легко використовувати для порівняння різних об'єктів одної або різної природи.

Проте існують і пару недоліків:

- за допомогою даного показника неможливо оцінити розміри втрат поза довірчим інтервалом;
- дана міра не є когерентною ризик-метрикою, оскільки вона не задовольняє властивість полуадитивності.

Ця міра часто використовується, оскільки її досить легко вичислити. Для цього існують велика кількість методів: параметричних і непараметричних. До них відносяться: історичний метод, параметричний

(відомий вид розподілу доходностей, часто беруть нормальний розподіл), метод Монте-Карло.

Крім вказаних вище метрик існує і багато інших. Наприклад: стрес-тестування, капітал під ризиком, показник максимальних втрат.

На даний момент практичні і наукові роботи постійно розвивають дані метрики, оскільки чим складніша система, що використовується в банках, тощо, тим більше ризиків і вони мають більший рівень.

### 1.9 Зашумованість даних

Чим більш розвинений світ, тим більше факторів, що впливають на формування певних систем і даних. Проте завжди треба пам'ятати, що крім показників, що справді мають вплив, існують фактори, що не мають логічного зв'язку з даними, але все-одно свій слід лишають. Наприклад, в будь-яких пристроях існує похибка вимірювань або обчислень, тому вимірюючи певний показник, туди можна додати і ця похибка, тим самим змінюючи справжні дані. Такі випадкові фактори називаються шумами, з якими необхідно боротись, оскільки шуми заважають оцінити дані як найближче до реальності.

Шум – це безладні коливання різної фізичної природи, що відрізняються складністю часової і спектральної структури.

В усіх фінансових даних присутні шуми, проте з ними рідко борються, оскільки це досить складна задача. Проте якщо припустити, що в даних присутній тільки білий шум, то задача стає набагато легшою. Білий шум – це шум, що підпорядковується нормальному закону. А відповідно з ним можна боротись. Для цього зазвичай використовується фільтрація даних.

## 1.10 Методи очищення даних від шумів

Необхідність очистки даних від шумів виникає щоразу як потрібно відділити реальні дані від шуму. Ціль процесу фільтрації даних – це якнайкраще відновити початковий сигнал на фоні перешкод, або визначення наявності корисного сигналу, або розділ кількох сигналів, що присутні у вхідній послідовності.

Всі види фільтрів можна поділити на класи: частотні, лінійні, нелінійні, комбіновані, гібридні, адаптивні.

В класі частотних фільтрів обробку проходять коефіцієнти розкладу зашумленого сигналу по базису Фур'є, або інших базисів. Такі алгоритми досить ефективні, вони не вимагають апіорної інформації, яка часто відсутня на практиці.

Лінійні методи часто використовуються в цифровій обробці, проте вони корисні лише при нормальному розподілі шумів.

В адаптивних фільтрах великі апертури використовуються в монотонних областях сигналу, а малі – недалеко від неоднорідностей, зберігаючи їх особливості.

Кожен метод має свої особливості і ситуації коли їх варто використовувати.

Проте на практиці найчастіше використовується фільтр Калмана для очистки шумів фінансових даних.

### 1.11 Висновки до розділу

В час розвитку інформаційних технологій дуже важливо приділяти увагу фінансовій системі, оскільки вона аналогічно вдосконалюється разом з математикою, технологіями, тощо.

На даний момент існує безліч людей, які займаються бізнесом. Досить часто їм необхідний додатковий капітал, яким їх можуть забезпечити продажі акцій компанії. Люди, що купують дані акції повинні мати гарантію, що гроші, які вони туди вкладуть, не пропали марно і принесли їм певний дохід.

Рідко коли компанії готові ділитися усіма своїми проблемами, тому потрібен якийсь інший механізм оцінки компанії і її акцій. Оскільки бездумна покупка акції несе за собою і певні ризики. В основному це ризики втрати всіх своїх вкладень. Для цього необхідно розробляти систему, яка б допомагала оцінювати ризики акцій.

В принципі існує багато різних підходів для оцінки фінансових ризиків, одною з найчастіше використовуваних є методологія VaR, Оскільки вона досить проста в обчисленні, і її легко інтерпретувати.

Також використовується волатильність, проте зазвичай не просто сам показник, а методи, що використовують волатильність.

Проте також не потрібно забувати, що в даних завжди присутній шум, з яким необхідно боротись, для отримання кращих результатів.

Постановка задачі:

1. Виконати огляд сучасних математичних інструментів і методів для оцінки ризику акцій.
2. Визначити методи для фільтрації шумів в даних.
3. Зібрати необхідні дані.
4. Визначити методологію, яка краще зможе виконувати поставлену задачу.

5. Розробити програмний продукт для оцінки ризику акцій.
6. Провести практичне застосування на забраних даних.
7. Виконати аналіз результатів.

## РОЗДІЛ 2. ОПИС МОДЕЛЕЙ ДЛЯ ОЦІНКИ РИЗИКУ АКЦІЙ

### 2.1 Фільтр Калмана і його модифікації

Фільтр Калмана – це ефективний рекурсивний фільтр, що оцінює вектор стану динамічної системи, використовуючи ряд неповних і зашумованих вимірів.

Фільтр Калмана і лінійно-квадратичний регулятор – можливе вирішення більшості фундаментальних задач в теорії керування.

Даний метод реалізований в часі і оперує оцінками не тільки стану, але й оцінками невизначеності вектора стану опираючись на формулу Байєса умовної ймовірності.

Даний фільтр працює з шумами будь-якого розподілу, проте коли шум має нормальний розподіл, то фільтр Калмана дає точну оцінку умовної ймовірності розподілу стану системи.

Класичний фільтр Калмана є рівнянням для розрахунку першого і другого моментів щільності ймовірності (математичне сподівання і дисперсія). Для нормального розподілу математичне сподівання і дисперсія задають повністю щільність розподілу. Тоді фільтр Калмана повністю описує вектор стану як випадкову векторну величину.

В цьому випадку оцінки математичного сподівання є оптимальними оцінками по критерію середньоквадратичної похибки. Даний момент обумовлює широке використання цього фільтру.

Розглянемо роботу даного фільтру:

Нехай є динамічна система:

$$x = F * x + w(Q), \quad (2.1)$$

де  $F$  – матриця переходів,



$w(Q)$  – випадковий процес (шум) з нульовим математичним сподіванням і матрицею коваріацій  $Q$ .

Тепер поспостерігаємо переходи станів системи з відомою похибкою вимірювань в кожен момент часу. Видалення шуму за допомогою фільтра Калмана складається з двох етапів – екстраполяція і корекція.

Задамо параметри системи:

$Q$  – матриця коваріацій шуму.

$H$  – матриця спостережень.

$R$  – коваріація шуму спостережень.

$P = Q$  – початкове значення коваріаційної матриці для вектора стану.

$z(t)$  – стан системи, що спостерігається.

$x = z(0)$  – початкове значення оцінки стану системи.

Для кожного спостереження знаходиться відфільтроване значення стану.

Крок 1. Екстраполяція.

1. Екстраполяція стану системи відбувається за формулою:

$$x = F * x, \quad (2.2)$$

2. Пошук коваріаційної матриці для екстрапольованого вектора стану:

$$P = F * P * F^T + Q, \quad (2.3)$$

Крок 2. Корекція.

3. Знаходимо вектор помилки:

$$y = z - H * x, \quad (2.4)$$

4. Знаходимо коваріаційну матрицю для вектора помилки:

$$S = H * P * H^T + R, \quad (2.5)$$

5. Знаходимо коефіцієнти підсилення Калмана:

$$K = P * H^T * S^{-1}, \quad (2.6)$$

6. Коректуємо оцінки вектора стану:

$$x = x + K * y, \quad (2.7)$$

7. Коректуємо коваріаційну матрицю оцінки вектора системи:

$$P = (I - K * H) * P, \quad (2.8)$$

Адаптивний фільтр Калмана – це модифікація фільтра Калмана.

Нехай існує система, що описується формулами:

$$x = F * x + w, \quad w(Q) \sim N(0, Q), \quad (2.9)$$

$$z = H * K + V, \quad V \sim N(0, R), \quad (2.10)$$

де  $x$  – вектор стану,

$z$  – спостережувана змінна,

$$X_0 \sim N(X_0, P_0),$$

$\theta = [X_0, P_0, F, H, Q, R]$  – невідомі. Необхідно визначити  $\theta$ .

Тоді прогноз вираховується за формулою:

$$x = F * x, \quad (2.11)$$

$$P = F * P * F^T + Q, \quad (2.12)$$

де  $x$  – новий стан системи,

$P$  – коваріаційна матриця стану системи.

Тепер проводимо корекцію.

Знаходимо помилку за формулою:

$$y = z - H * x, \quad (2.13)$$

$$S = H * P * H^T + R, \quad (2.14)$$

де  $y$  – вектор помилки,

$S$  – коваріаційна матриця для вектора помилки.

Далі шукається оптимальний коефіцієнт підсилення Калмана:

$$K = P * H^T * S^{-1}. \quad (2.15)$$

В кінці відбувається коригування стану системи:

$$x = x + K * y, \quad (2.16)$$

$$P = (I - K * H) * P. \quad (2.17)$$

І нарешті згладжування за допомогою фільтра Калмана:

$$L_k = P_{k|k} F_k^T P_{k+1|k}^{-1}, \quad (2.18)$$

$$\hat{X}_{k|N} = \hat{X}_{k|k} + L_k (\hat{X}_{k+1|N} - \hat{X}_{k+1|k}), \quad (2.19)$$

$$P_{k|N} = P_{k|k} + L_k (P_{k+1|N} - P_{k+1|k}) L_k^T. \quad (2.20)$$

Маючи оцінку  $\theta$  на попередньому кроці:

$$\tilde{Y}_k(\theta) = Z_{k+1} - H \hat{X}_{k+1|k}, k = 1, 2, \dots, N, \quad (2.21)$$

$$S_{k|k-1}(\theta) = H P_{k|k-1} H^T + R. \quad (2.22)$$

Наступний етап – це побудова функції правдоподібності для вектора шуму:

$$L(\theta) = \prod_{k=1}^N \frac{2}{(2\pi)^{\frac{n}{2}}} |S_{k|k-1}(\theta)|^{-\frac{1}{2}} \exp \left( -\frac{1}{2} \tilde{Y}_k(\theta)^T S_{k|k-1}^{-1} \tilde{Y}_k(\theta) \right). \quad (2.23)$$

Розв'язавши задачу максимізації для функції правдоподібності знаходимо нову оцінку параметра  $\theta$ .

Тепер можна записати алгоритм для адаптивного фільтра Калмана:

$$X_0 = \hat{X}_{0|N}. \quad (2.24)$$

$$H = \sum_{k=0}^N Z_k \hat{X}_{k|N}^T \left( \sum_{k=0}^N (P_{k|N} + \hat{X}_{k|N} \hat{X}_{k|N}^T) \right)^{-1}. \quad (2.25)$$

$$F = \left( \sum_{k=1}^N (P_{k|N} + \hat{X}_{k|N} \hat{X}_{k-1|N}^T) \right) \left( \sum_{k=1}^N (P_{k-1|N} + \hat{X}_{k-1|N} \hat{X}_{k-1|N}^T) \right)^{-1}. \quad (2.26)$$

$$P_0 = P_{0|N}. \quad (2.27)$$

$$R = \frac{1}{N+1} \left( \sum_{k=0}^N Z_k Z_k^T - \sum_{k=0}^N Z_k \hat{X}_{k|N}^T \left( \sum_{k=0}^N (P_{k|N} + \hat{X}_{k|N} \hat{X}_{k|N}^T) \right)^{-1} \hat{X}_{k|N} Z_k^T \right) \quad (2.28)$$

$$Q = \frac{1}{N} \left( \sum_{k=0}^N (P_{k|N} + \hat{X}_{k|N} \hat{X}_{k|N}^T) - \sum_{k=1}^N (P_{k-1|N} + \hat{X}_{k-1|N} \hat{X}_{k-1|N}^T) * \right. \\ \left. * \left( \sum_{k=0}^N (P_{k-1|N} + \hat{X}_{k-1|N} \hat{X}_{k-1|N}^T) \right)^{-1} \sum_{k=1}^N (P_{k|N} + \hat{X}_{k|N} \hat{X}_{k|N}^T) \right) \quad (2.29)$$

1. Встановлюємо початкове значення параметра  $\theta^{(0)}, i = 0$ .
2. Знаходимо вектор помилки і коваріаційну матрицю до неї.
3. Обчислюємо функцію правдоподібності для помилки.
4. Знаходимо оцінку параметра  $\theta$  за допомогою задачі максимізації функції правдоподібності.
5. Отримуємо новий вектор параметрів  $\theta^{(i+1)}, i = i + 1$ .
6. Повторюємо пункти 2-5 поки не буде збіжності.

## 2.2 Моделі для оцінювання часових рядів

Будь-які системи, які описуються часовими рядами можуть бути стаціонарними і нестаціонарними. Стаціонарність означає незалежність

характеристик (математичне сподівання, дисперсія) ряду від часу. Формально це визначається так:

$$\begin{aligned} E(x_t) &= \text{const} \text{ (математичне сподівання)} \\ D(x_t) &= E((x_t - E(x_t))^2) = \text{const} \text{ (дисперсія)} \\ \text{Cov}(x_t) &= E((x_t - E(x_t))(x_{t-s} - E(x_{t-s}))) = \text{const} \text{ (кореляція)} \end{aligned} \quad (2.30)$$

Процес стаціонарний при умові, що виконуються усі три умови. Якщо ж хоча б одна не виконується, то ряд є нестаціонарним, тобто в процесі може бути присутніми тренд, сезонність або періодичність. Це ускладнює роботу з такими процесами.

Кожен тип нестаціонарності має свої особливості.

Якщо ж говорити про фінансові процеси, то мірою волатильності фондових процесів часто береться дисперсія. Оскільки в таких процесах волатильність не стійка у часі, то фінансові процеси є гетероскедастичними (нестаціонарними зі змінною дисперсією у часі). Це означає, що розкид кожної випадкової величини навколо математичного сподівання змінюється у часі.

Одним із найпростіших способів описання змінної дисперсії є модель типу AR(q) для оцінки квадратів залишків. Тобто:

$$\hat{\varepsilon}^2(k) = a_0 + \sum_{i=1}^q a_i \hat{\varepsilon}^2(k-i) + v(k), \quad (2.31)$$

де  $v(k)$  – білий шум з нульовим середнім.

Отримане рівняння є авторегресійним умовно гетероскедастичним рівнянням (ARCH).

Дане рівняння можна використовувати для оцінки умовної дисперсії:

$$E[\hat{\varepsilon}^2(k+1)] = a_0 + \sum_{i=1}^q a_i \hat{\varepsilon}^2(k-i) + v(k). \quad (2.32)$$

Залишки, що використовуються в моделі можна знайти за допомогою інших моделей: авторегресії, тощо.

В попередній моделі збурення впливало на процес адитивно, проте вплив може бути і мультиплікативним. Більше того, ми цього майже ніколи не знаємо наперед.

В цьому випадку модель буде мати вигляд:

$$\hat{\varepsilon}^2(k) = v^2(k) \left( a_0 + \sum_{i=1}^q a_i \hat{\varepsilon}^2(k-i) \right)^{1/2}, \quad (2.33)$$

де  $v(k)$  – білий шум з у мультиплікативній формі. Причому  $v(k) \sim N(0, 1)$ , змінні  $\varepsilon(k-i)$  і  $v(k)$  – незалежні.

Модель GARCH є розширенням моделі ARCH. Вона говорить, що умовна дисперсія може бути процесом АРКС. Нехай:

$$\varepsilon(k) = v(k)(h(k))^{1/2}, \quad (2.34)$$

де  $\sigma_v^2 = 1$ , а умовна дисперсія оцінюється за формулою (модель АРКС(p, q)):

$$h(k) = a_0 + \sum_{i=1}^q a_i \varepsilon^2(k-i) + \sum_{i=1}^p \beta_i h(k-i). \quad (2.35)$$

Оскільки  $v(k)$  – білий шум і не корелює з  $\varepsilon(k-1)$ , то математичне сподівання як умовне, так і безумовне для помилки становить нуль.

Знаючи усе це можна знайти вибірккові умовне математичне сподівання і умовну дисперсію:

$$\underline{y}_B(k) = \frac{1}{k} \sum_{i=1}^k y(i), \quad (2.36)$$

$$h_B(k) = \frac{1}{k-1} \sum_{i=1}^k (y(i) - \underline{y}_B(k))^2, \quad k = 2, 3, \dots, N. \quad (2.37)$$

Для забезпечення скінченності умовної дисперсії необхідно, щоб корені характеристичного рівняння знаходились в серединні одиничного кола.

Для даної моделі значущий лаг знаходимо не по АКФ, а за допомогою Q-статистики Лjung-Бокса:

$$Q = N(N+2) \sum_{i=1}^n \frac{\rho(s)}{(N-s)}. \quad (2.38)$$

Інтерпретувати цю статистику можна таким чином: якщо  $\hat{\varepsilon}^2(k)$  некорельовані між собою, тоді статистика має розподіл  $\chi^2$  з  $n$  ступенями свободи. Тоді за нульову гіпотезу береться:  $\hat{\varepsilon}^2(k)$  некорельовані між собою. Якщо ми відмовились від нульової гіпотези, то існує автокореляція до порядку  $n$  у часовому ряді.



В цих моделях також існує багато модифікацій.

Якщо ж модель не гетероскедастична можна використовувати моделі АР або АРКС.

Авторегресійна модель (АР) дозволяє знаходити певних показник через інші змінні у вигляді їх комбінацій. Суть даного методу в представленні значення часового ряду в один момент через значення в попередні моменти часу.

Загальна випадок методу авторегресії описується за формулою:

$$x_t = c + \sum_{i=1}^p a_i x_{t-i} + \varepsilon_t, \quad (2.39)$$

де  $x_{t-i}$  – значення часового ряду в момент  $t - i$ ;

$c$  – постійне значення;

$a_1, a_2, \dots, a_p$  – параметри моделі, коефіцієнти авторегресії;

$\varepsilon_t$  – білий шум;

$x_t$  – значення, яке прогнозується за допомогою моделі (1.7) на момент часу  $t$ ;

$p$  – порядок авторегресійної моделі.

Дуже часто вільний коефіцієнт беруть рівним нулю, щоб спростити модель.

Основна задача в цьому методі визначити самий порядок моделі  $p$ . Робиться це за допомогою часткової автокореляційної функції. Дана функція є характеристикою часового ряду, що допомагає показати залежність між різними членами ряду.

На практиці дуже рідко вдається визначити порядок моделі, тому за допомогою ЧАКФ визначається максимальне значення порядку. І перебираються усі моделі з порядком  $p$  і нижче. Тоді вибирається модель, яка найкраще описує процес.

Метод авторегресії з ковзним середнім (АРКС) – це певна комбінація методів АР і КС. В даному методі ми починаємо враховувати білий шум, що теж змінюється у часі. Дану модель можна представити в такому вигляді:

$$x_t = c + \sum_{i=1}^p a_i x_{t-i} + \sum_{j=1}^q b_j \varepsilon_{t-j} + \varepsilon_t, \quad (2.40)$$

де  $x_{t-i}$  – значення часового ряду в момент  $t - i$ ;

$c$  – постійне значення;

$a_1, a_2, \dots, a_p$  – параметри моделі, коефіцієнти авторегресії;

$b_1, b_2, \dots, b_q$  – параметри моделі, коефіцієнти ковзного середнього;

$x_t$  – значення, яке прогнозується за допомогою моделі (1.8) на момент часу  $t$ ;

$\varepsilon_t$  – білий шум;

$p$  – порядок авторегресійної моделі;

$q$  – порядок моделі ковзного середнього.

Причому вважається, що шум має незалежні між собою члени ряду, а також розподілений за нормальним розподілом, де математичне сподівання дорівнює нулю, а дисперсія – 1.

Тут також необхідно знайти порядки АР і КС. Вони знаходяться за ЧАКФ і АКФ відповідно.

## 2.3 Нейронні мережі

Нейронні мережі – обчислювальні системи, натхненні біологічними нейронними мережами, що складають мозок тварин [23].

Дані системи мають основну особливість: вони постійно навчаються, тобто поступово покращують результати, на задачах, при цьому не потребують спеціальної розробки під задачу. Нейронні мережі використовуються у великій кількості задач. Одним таким прикладом є розпізнавання зображень, де система може навчитись розрізняти собаку і kota як різні об'єкти.

На Рисунку 2.1 зображено схему найпростішої стандартної нейронної мережі, де кружки це нейрони, а стрілки вказують на рух імпульсу від вихідного до вхідного нейрону.

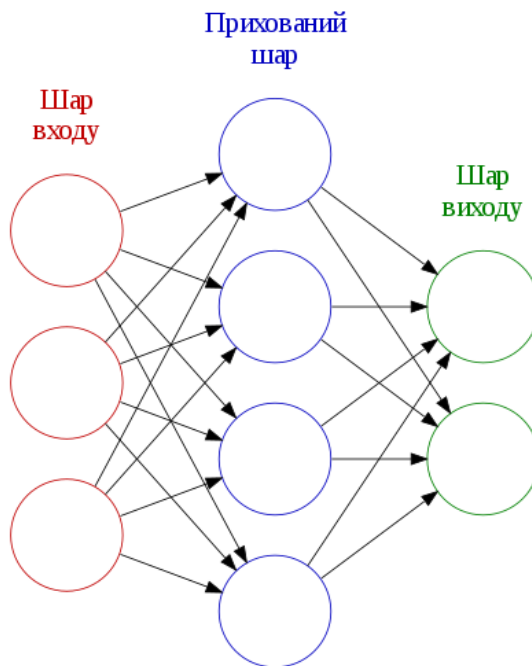


Рисунок 2.1 – Схема елементарної нейронної мережі

Нейронні мережі складаються з наступних елементів:

1. Нейрони. Він включає в себе: збудження або активацію нейрона, функцію збудження (незмінна поки функція навчання не вирішить її замінити), поріг (за необхідністю), функцію виходу (часто буває просто тотожною функцією).

2. З'єднання та ваги. В мережі між нейронами існують зв'язки, які передають інформацію (вихід нейрона) до наступного нейрона. Таким зв'язкам присуджуються ваги.

3. Функція поширення. Дана функція часто має такий вигляд:

$$p_j(t) = \sum_i a_i(t)w_{ij}, \quad (2.41)$$

де  $p_j$  – вхід нейрона  $j$ ,

$a_i$  – виходи попередніх нейронів.

4. Правило навчання. Відповідає за змінну параметрів мережі, щоб вхід був зв'язаний з виходом з мережі правильним чином.

Нейронні мережі бувають різних видів. Кожен вид характеризується власними особливостями.

Згорткові нейронні мережі – це мережі прямого поширення, які включають в себе згорткові внутрішні шари.

Метод групового урахування аргументів. Даний метод надає змогу використовувати автоматизовану оптимізацію моделей.

Дуже популярними в прогнозування часових рядів є метод довгої короткочасної пам'яті або рекурентна нейронна мережа. Даний метод дозволяє не відкидати складову часу з задачі, а навпаки, будуються з врахуванням часу і залежності змінних в часі. Що є дуже зручним для їх використання при аналізі і прогнозуванні часових рядів.

При використанні такої мережі на вхід подається вектор даних на певний момент часу, на наступний вхід вектор на наступний момент часу і т.д. При цьому мережа навчаючись враховує не тільки цей вхід і прораховані для них параметри, але й усі попередні моменти часу, тим самим даючи змогу даним в майбутній момент часу бути залежним від попередніх моментів часу.

На кожен момент часу дані проходять всі внутрішні шари мережі, але процес не закінчується на виході, а додається до нових вхідних даних, тим самим уточнюючи модель.

Дана особливість доступна не для всіх методів машинного навчання, тому є дуже застосованою.

Схема такої мережі зображено на Рисунку 2.2.

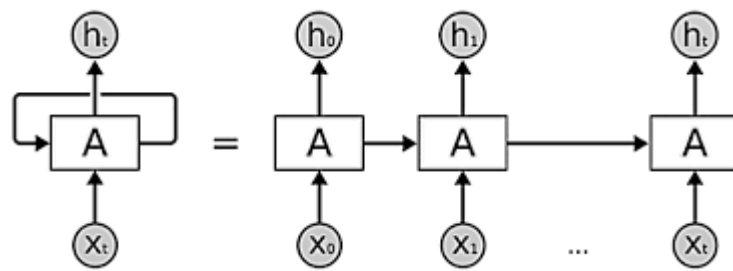


Рисунок 2.2 – Схема елементарної рекурентної нейронної мережі

Для рекурентних нейронних мереж часто використовується алгоритм LSTM.

Дана мережа має вигляд проілюстрований на Рисунку 2.3.

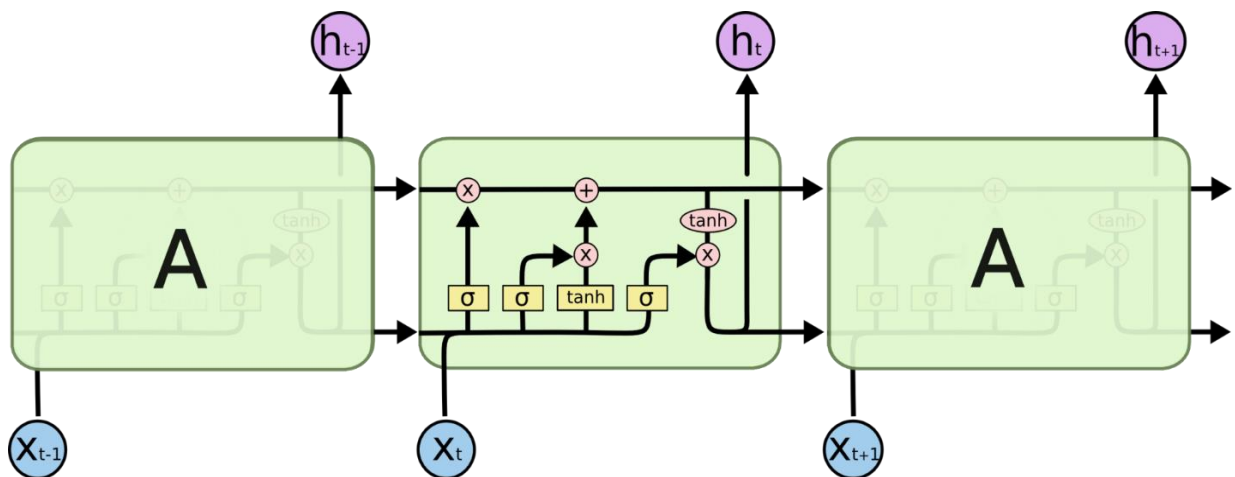


Рисунок 2.3 – Схема LSTM

На схемі (Рисунок 2.3) видно, що мережа включає в себе чотири внутрішні шари, які взаємодіють між собою. Зазвичай це шари, що відповідають сігмоїді і тангенсу гіперболічному (функції активації).

Першим кроком в цій моделі є визначення того, яку інформацію необхідно викинути з стану комірки. Такий шар називається сигмоїдним шаром. Він дивиться на вхідні дані і дані з попередніх кроків і присуджує 0 або 1 відповідно до формули сигмоїди.

Функція активації має вигляд:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f). \quad (2.42)$$

На наступному кроці вирішується, яка інформація буде далі зберігатись в стані комірки, а яке значення варто оновити.

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i), \quad (2.43)$$

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c). \quad (2.44)$$

Третій шар відповідає за заміну станів на нові. Це відбувається за формулою:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t. \quad (2.45)$$

В кінці вирішується, яку саме інформацію необхідно отримати на виході. Для цього використовуються наступні формули:

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o), \quad (2.46)$$

$$\text{де } h_t = o_t * \tanh(C_t). \quad (2.47)$$

На цьому кроці вибирається, що саме потрібно вивести, потім обране проходить певну трансформацію і відправляється на вихід.

## 2.4 Перевірка наявності гетероскедастичності

Наявність гетероскедастичності в процесах може привести до великих похибок в оцінці таких процесів більшістю методів. Тому є дуже важливим проводити перевірку на гетероскедастичність.

Найпростішим способом перевірки є аналіз графіків залишків регресії. Проте це не завжди ефективно. Тому існує ряд строгих тестів для перевірки даної властивості.

Тест Уайта. Нехай є лінійна регресія. Тест використовує залишки регресії, що оцінюються за методом найменших квадратів.

Нульовою гіпотезою тут буде: відсутність гетероскедастичності. Для даної перевірки використовується LM-статистика.

$$LM = nR^2, \quad (2.48)$$

де  $R^2$  – коефіцієнт детермінації допоміжної регресії,

$n$  – кількість спостережень допоміжної регресії.

При відсутності гетероскедастичності ця статистика має асимптотичний розподіл  $\chi^2$  з  $N - 1$  ступенями свободи, де  $N$  – кількість параметрів допоміжної регресії.

Отже, якщо значення статистики більше критичного значення, то нульова гіпотеза не приймається, в іншому випадку вважається, що процес гомоскедастичний.

Тест Голдфелда-Куандта. Використовується у випадку, якщо є припущення щодо того, що стандартне відхилення похибки може бути пропорційне певній змінній. Також похибки повинні бути нормально розподіленими випадковими величинами.

Спершу йде впорядкування даних по спаданню незалежної змінної, щодо якої є підозри на гетероскедастичність.

За МНК оцінюється регресійна модель для двох різних вибірок – перших і останніх  $m$  спостережень в упорядкуванні. Для цих моделей знаходяться суми квадратів залишків і розраховується F-статистика:

$$F = \frac{RSS_1}{RSS_2}. \quad (2.49)$$

Ця статистика при відсутності гетероскедастичності має розподіл Фішера  $(m - k, m - k)$  ступенями свободи. Отже, якщо статистика більша за значення даного розподілу при даному рівні значимості, то нульова гіпотеза не приймається, тобто є гетероскедастичність. В іншому випадку вважається, що присутня гомоскедастичність.

Крім того можна зробити перевірку за допомогою P-значення. Якщо  $P(F) < \alpha$ , де  $\alpha$  – рівень значимості, то гетероскедастичність значима, в іншому випадку – ні.

Тест рангової кореляції Спірмена – це непараметричний статистичний тест, що дозволяє перевірити гетероскедастичність випадкових похибок регресійної моделі. Особливість даного тесту в тому, що не має конкретизації форми можливої залежності дисперсії випадкових похибок моделі від тої чи іншої змінної.

За допомогою МНК оцінюється лінійна регресійна модель і визначаються залишки моделі.



Потім проходить ранжування по залишкам, і визначається коефіцієнт рангової кореляції Спірмена:

$$\hat{\rho} = 1 - \frac{6 \sum d_t^2}{n(n^2 - 1)}, \quad (2.50)$$

де  $d_t^2$  – квадрати різниці рангів змінних похибки і змінної, щодо якої передбачається залежність.

При справедливості нульової гіпотези (відсутність гетероскедастичності) статистика  $\hat{\rho}\sqrt{n-1}$  асимптотично має стандартний нормальний розподіл. Відповідно, якщо значення статистики більше за критичне значення цього розподілу, то гетероскедастичність приймається значимою, в іншому випадку гетероскедастичність незначима і приймається гомоскедастичність.

## 2.5 Методи перевірки моделі на адекватність

Після того як модель побудована, її необхідно перевірити, чи дійсно вона гарно описує наші дані. Чи можна взагалі використовувати цю модель для оцінки даних.

Модель є адекватною, якщо рівні кінцевої послідовності коливаються випадково, залишкова компонента розподілена за нормальним законом, математичне сподівання залишкової компоненти дорівнює 0 і рівні залишкової компоненти не корелюються між собою.

Для перевірки даних властивостей існує безліч критеріїв.

Зазвичай перевірка відбувається на основі залишків.

RS-критерій.

Це критерій для дослідження розподілу залишкової компоненти.

Нульова гіпотеза: залишкова компонента розподілена за нормальним законом.

Критерій розраховується за формулою:

$$RS = \frac{(\varepsilon_{max} - \varepsilon_{min})}{\sigma}, \quad (2.51)$$

де  $RS$  – значення критерію;

$\varepsilon_{max}$  – максимальне значення похибки моделі;

$\varepsilon_{min}$  – мінімальне значення похибки моделі;

$\sigma$  – середньоквадратичне відхилення залишків:

$$\sigma = \sqrt{\frac{\sum_{i=0}^n (\varepsilon_t - \underline{\varepsilon}_t)^2}{n - 1}}. \quad (2.52)$$

Знайшовши значення критерію необхідно провести порівняння з табличними значеннями для даного критерію. Якщо значення критерію входять в заданий інтервал, то з відповідним рівнем значимості приймається нульова гіпотеза, в іншому випадку гіпотеза не приймається і модель вважається неадекватною.

Log likelihood.

Ще один критерій для перевірки нормального розподілу залишків моделі. Нульова гіпотеза приймається у випадку, якщо логарифмічна функція правдоподібності зменшується з часом.

Статистика Дурбіна-Ватсона.

Статистика Дурбіна-Ватсона допомагає перевірити корельованість між собою залишків моделі. Коефіцієнт розраховується за наступною формулою:

$$DW = \frac{\sum_{t=2}^N (\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^N \varepsilon_t^2}, \quad (2.53)$$

де  $\varepsilon_t$  – похибка моделі в момент часу  $t$ ;

$DW$  – коефіцієнт Дурбіна-Ватсона.

Значення коефіцієнт Дурбіна-Ватсона завжди лежить у відрізку  $[0; 4]$ .

Якщо коефіцієнт  $DW$  збігається до 2 коли час прямує до безкінечності, тоді вважають. Що залишки моделі некорельовані між собою. В іншому випадку, залишки корельовані.

Критерій Айкаке.

Даний критерій служить для перевірки адекватності моделі. Спершу знаходиться значення критерію за формулою:

$$AIC = \frac{2l}{T} + \frac{2k}{T}, \quad (2.54)$$

де  $AIC$  – значення самого критерію;

$l$  – значення логарифмічної функції правдоподібності (була приведена вище).

Для найкращої моделі цей показник має бути мінімальним.

## 2.6 Критерії якості оцінок прогнозів

Одним з основних етапів побудови прогнозової моделі є оцінка якості побудованих прогнозів. За допомогою цих оцінок можна зрозуміти, що

модель, яка була адекватною, не досить гарно описує модель. В цьому випадку необхідно ввести зміни і корегування моделі.

До оцінок якості прогнозів використовують такі характеристики: надійність, точність, достовірність, похибки прогнозів.

Надійність прогнозів – це міра невизначеності поведінки моделі з часом.

Достовірність прогнозів – це ймовірність справдження даного прогнозу для заданого варіанта або довірчого інтервалу.

Точність прогнозів – інтервальний розкид прогнозних траєкторій при фіксованому рівні достовірності.

На жаль, описати такі характеристики як надійність, точність, достовірність і визначити похибки прогнозів апріорі неможливо. Тому єдиний вихід розібрати фактори, що впливають на показники якості прогнозу:

- якість вхідної інформації;
- прогнозуюча модель;
- метод прогнозування;
- якість прогнозування.

Найбільші похибки виникають в результаті невдалого вибору метода прогнозування. В цьому випадку потрібно гарно розуміти структуру процесу, щоб підібрати хороший метод.

Багато похибок присутні і при побудові самих моделей. В цьому випадку можна використовувати значення похибок моделювання, а саме:

Абсолютна похибка моделі:

$$\Delta = |y - \hat{y}|, \quad (2.55)$$

де  $y$  – справжнє значення показника,

$\hat{y}$  – прогнозне значення отримане за допомогою моделі.

Середньоквадратична похибка:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}, \quad (2.56)$$

де  $y$  – справжнє значення показника,

$\hat{y}$  – прогнозне значення отримане за допомогою моделі,

$n$  – розмір вибірки.

Коефіцієнт детермінації.

Коефіцієнт детермінації – це доля дисперсії залежної змінної, що пояснюється використаною моделлю.

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}, \quad (2.57)$$

де  $SS_{res}$  – сума квадратів залишків регресії,

$SS_{tot}$  – загальна сума квадратів.

$$SS_{res} = \sum_{i=1}^n (y_i - \hat{y}_i)^2, \quad (2.58)$$

де  $y$  – справжнє значення змінної,

$\hat{y}$  – прогнозне значення отримане за допомогою моделі,

$n$  – розмір вибірки.

$$SS_{tot} = \sum_{i=1}^n (y_i - \underline{y})^2, \quad (2.59)$$

$$\underline{y} = \frac{1}{n} \sum_{i=1}^n y_i. \quad (2.60)$$

## 2.7 Висновки до розділу

В даному розділі було розглянуто методи фільтрації даних. Методи фільтрації дозволяють очистити дані від шуму, тобто викинути все лишнє з даних, щоб залишились тільки максимально чисті і справжні дані. Оскільки значень цін акцій, як і будь-які інші фінансові дані, включають в своє формування дуже багато факторів, в тому числі і фактори, що накопичують в собі теж зайві дані. Тому використання фільтрів описаних в розділі є хорошою ідеєю перед використанням методів прогнозування.

Для оцінок ризику акцій часто використовується волатильність. Вона зазвичай виступає як дисперсія. Оскільки зміна дисперсії в часі приводить до нестационарності процесу необхідно використовувати математичні моделі, які б дозволили гарно і адекватно описувати поведінку таких процесів. Що допоможе краще прогнозувати ціни акцій і оцінювати їх ризики.

Саме тому в розділі було розглянуто математичні методи моделювання гетероскедастичних і гомоскедастичних процесів. А також рекурентні нейронні мережі, які теж можуть бути корисними для такого типу задач.

## РОЗДІЛ 3. АНАЛІЗ ОБЧИСЛЮВАЛЬНИХ ЕКСПЕРИМЕНТІВ ТА ОГЛЯД ПРОГРАМНОГО ПРОДУКТУ

### 3.1 Вхідні дані

В даній роботі використовувались дані про ціни акцій компанії Google за 2010 – 2016 роки.

В першу чергу було проведено візуальний аналіз часового ряду. Для цього представлено графіки часового ряду, автокореляційної функції і частково-автокореляційної функції на Рисунку. 3.1.

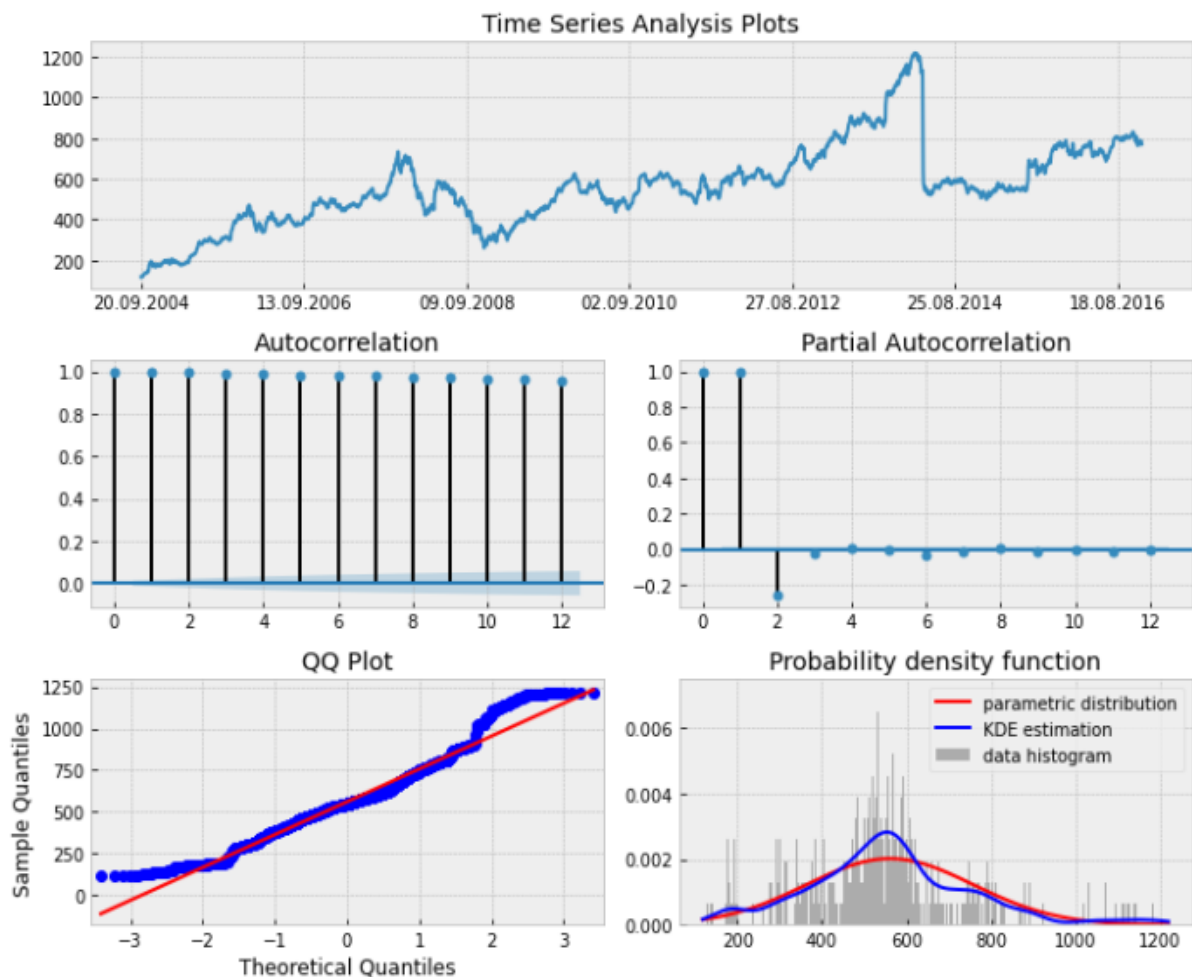


Рисунок 3.1 – Графік значень цін акцій компанії Google

За графіком частково-автокореляційної функції очевидно, що параметр для авторегресії варто брати рівний 3. А по кореляційній функції можна зробити висновок, що потрібно брати ковзне середнє з параметром 12.

По графіку щільності зображеному на Рис. 3.1 можна зазначити, що розподіл цін схожий на нормальний, проте являється трохи вищим за нормальний розподіл, що може вказувати на розподіл Стюдента. Проте він також має витягнутий правий хвіст.

Висновок про критерій Стюдента підтверджують наступні характеристики часового ряду представлені в Таблиці 3.1.

Таблиця 3.1 – Описова статистика по часовому ряду

Показник	Значення
Min	117.84
Max	1220.17
Mean	560.577
Variance	38977.084
Skewness	0.572
Kurtosis	1.008278

### 3.2 Результати оцінювання

Для розрахунку показника VaR спершу необхідно знайти і спрогнозувати волатильність ряду. Для цього використовувались кілька методик.



Спершу було застосовано фільтр Калмана для очистки даних від випадкових шумів, що дало змогу покращити точність прогнозування волатильності, а отже, і збільшити точність оцінки ризику.

Також було проведено тест Engle для перевірки гетероскедастичності залишків моделі. Ймовірність отримана в результаті дорівнює 0, що менше за поріг 0.05. Отже, можна відхилити нульову гіпотезу про незмінність дисперсії в часі.

Наступним кроком було підібрати модель, яка б найкраще могла оцінити волатильність.

В першу чергу перевірялись моделі типу ARMA. Для цього необхідно було визначити параметри моделей.

На Рисунку 3.2. представлено значення АКФ і ЧАКФ для квадратів залишків моделі, які допоможуть визначити значення параметрів для моделей.

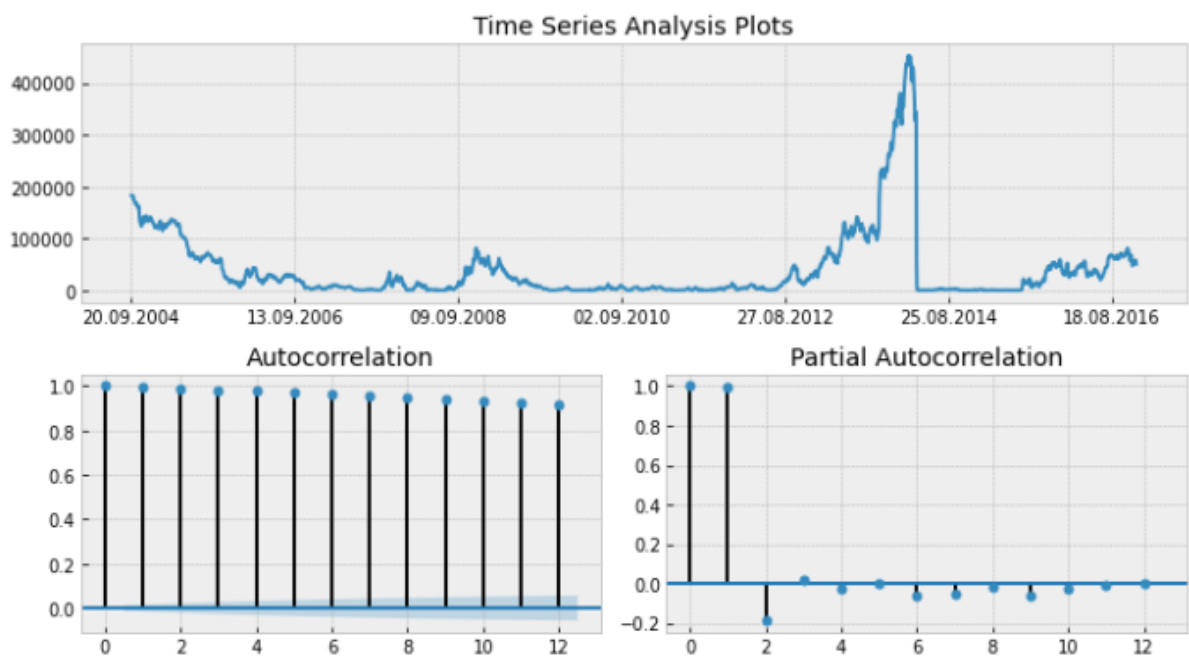


Рисунок 3.2 – Значення АКФ і ЧАКФ для квадратів залишків моделі

Судячи зі значень АКФ і ЧАКФ параметри можуть бути більше 1, тому порівнюємо модель з різними параметрами, щоб оцінити найкращу.

Таблиця 3.2 – Оцінки моделі GARCH

Модель	MAE	MAPE	THEIL
GARCH(1, 1)	46.703393537471	13.60310954523587	0.35711302848347 1
GARCH(2, 2)	46.703398699012	13.60310685149867	0.35971130284802
GARCH(2, 8)	46.702794762397	13.51956147319662	0.35971123179284 8
GARCH(3, 12)	46.059252938935	11.73202941090763	0.35708675806683 4

Судячи по Таблиці 3.2 моделі дають цілком прийнятні прогнози. В ході експерименту моделі EGARCH і TARARCH показали себе на такому ж рівні. Була спроба використати не нормальний розподіл, а розподіл Стюдента, він давав такі ж результати.

Для прогнозування часових рядів можна використати ще один метод: рекурентні нейронні мережі.

В ході написання магістерської дисертації було розроблено рекурентну нейронну мережу, яка працювала по принципу системи довгої короткочасної пам'яті.

Данна нейронна мережа отримувала на вхід справжню волатильність. В самій мережі присутні два внутрішніх шара. Перший шар – це LSTM, другий шар – Dense.

В таблиці 3.3 представлено результати роботи даного методу.

Таблиця 3.3 – Результати роботи рекурентної нейронної мережі на тренувальних і тестових даних

Параметр	Значення
rmse	13.20
rmse_test	7.25
mae	5.383441355970177
mae_test	5.149382323272167
mape	0.12988149290623077
mape_test	0.16886881767284728
theil	0.028672472752050742
theil_test	0.02116020195068419

В таблиці 3.4 представлено порівняння результатів отриманих по нейронній мережі і по моделі GARCH(1, 1). Було обрано саме модель GARCH(1, 1) серед інших моделей такого типу, оскільки інші методи не дають сильних покращень, тобто вони всі дають приблизно однакові результати.

Таблиця 3.4 – Порівняння результатів моделей

Модель	MAE	MAPE	THEIL
GARCH(1, 1)	46.703393537471	13.60310954523587	0.357113028483471
RHM	5.383441355970	0.129881492906230	0.0286724727520507

По результатам таблиці 3.4 можна зробити висновок, що рекурентна нейронна мережа має кращі характеристики, а отже, і дає кращі результати.

Отже, для подальшого обчислення показника VaR використовуються прогнози, отриманими методом RHM.

Після прогнозування волатильності можна розрахувати параметр Value-at-Risk. Даний показник можна розрахувати знаючи розподіл часового ряду, і якщо він є нормальним або логнормальним, тоді розраховується по спеціальним формулах. Оскільки на початку пункту був зроблений висновок про схожість розподілу до нормального і до розподілу Стюдента, то розрахуємо показник для обох випадків. Результати представлені в таблиці 3.5.

Таблиця 3.5 – Значення VaR для акцій

Показник	Значення
VaR 95%	493.738
VaR-t 95%	493.743
VaR 99%	466.099
VaR-t 99%	466.042

### 3.3 Функціональна схема алгоритму

В ході роботи була розроблена і реалізована програма, що призначена для аналізу ризику акцій фінансового ринку. Вона базується на прогнозуванні волатильності методом рекурентних нейронних мереж і оцінюванні ринкового ризику за методологією VaR, дані розрахунки робляться на основі прогнозів волатильності.

На Рисунку 3.3 представлено схему роботи алгоритму.



Рисунок 3.3 – Блок-схема роботи алгоритму

### 3.4 Висновки до розділу

Розроблено СППР, яка допомагає приймати рішення при купівлі чи продажі акцій. Дана система прогнозує волатильність і на її основі обчислює значення ризику VaR, тобто втрати, що не перевищать значення ризику з заданою ймовірністю. В роботі використовувалось значення ймовірностей 95% і 99%.

В роботі було представлено аналіз моделей, які часто використовуються для таких задач. Проте як було описано в розділі методи типу ARMA дуже погано справлялись з даними і давали дуже неточні

прогнози. За рахунок цього було прийнято рішення розробити рекурентну нейронну мережу для даної задачі. Дана мережа гарно справляється з задачею, має досить хорошу точність на відмінну від інших методів.

На основі даної рекурентної мережі було розроблено систему підтримки прийняття рішень для аналізу ризиків акцій.

Даний алгоритм повністю задовольняє основні вимоги до СППР: легкий у використанні, видає однозначні результати: значення VaR, на основі яких користувач приймає рішення про купівлю чи не купівлю акцій.

## РОЗДІЛ 4. РОЗРОБКА СТАРТАП ПРОЕКТУ

### 4.1 Постановка задачі стартап-проекту

Своєчасне оцінювання ризику необхідне на усіх етапах прийняття рішень в багатьох сферах діяльності. Принципи ризик-менеджменту однакові для всіх сфер діяльності незважаючи на принципову відмінність напрямків. Мету ризик-менеджменту можна визначити таким чином: визначення характеру протікання процесу протягом усього історичного періоду та прогнозування його поведінки у майбутньому, ідентифікацію основних ризикових компонент та оцінювання їх впливу на подальшу діяльність установи.

Вагомі відмінності починаються на етапі підбору методів аналізу та математичного апарату. Тому для отримання хороших результатів в сфері аналізу відокремлюють системи для різних напрямків. Тому для оцінки ринкових ризиків розробляються окрема система. Дана система дає можливість завантажити дані в зручному форматі, швидко проаналізувати взаємозв'язки та спрогнозувати дані, оцінити ризики для заданого рівня довіри. Завдяки такій системі можна вчасно виявити можливі втрати та своєчасно зреагувати на них.

Для розробки стартап-проекту та виведення його на ринок необхідно провести детальне дослідження, яке передбачає виконання наведених нижче чотирьох кроків.

Перший крок. Маркетинговий аналіз стартап-проекту:

- розробка опису ідеї проекту, визначення основних напрямків використання товару чи послуги, формування основних відмінностей від конкурентів;
- аналіз ринкових можливостей;
- розробка стратегії для виведення товару на ринок.

Другий крок. Організація стартап-проекту, що включає:

- календарний план реалізації та запуску продукту;
- плановий обсяг виробництва, та відповідно матеріальних ресурсів і персоналу;
- витрати для реалізації проекту та запуску його.

Третій етап. Фінансово-економічний аналіз та аналіз ризиків:

- обсяг інвестиційних витрат;
- розрахунок фінансових показників проекту та визначення інвестиційної привабливості проекту;
- визначення основних ризиків проекту.

Четвертий етап. Комерціалізація проекту:

- визначення цільової групи інвесторів;
- складання інвестиційної пропозиції;
- визначення основних каналів для просування oferty інвесторам.

## 4.2 Карта проекту

Проект полягає у створенні системи прийняття рішень для автоматичного оцінювання ринкових ризиків фінансового ринку. Дана програма аналізує дані введені дані і видає оцінку ризиків.

В таблиці 4.1 представлено основну інформацію по проекту.



Таблиця 4.1 – Інформаційна карта проекту

Назва проекту	Система аналізу ринкових ризиків на базі методів машинного навчання.
Автори проекту	Пущик Оксана
Коротка анотація	Система дозволяє оцінювати ризики, на основі яких користувач приймає рішення спрямоване на ріст прибутку.
Термін реалізації проекту	12 місяців
Необхідні ресурси	Інтелектуальні: спеціаліст з аналізу даних; розробник; тестувальник; продуктовий менеджер проекту. Матеріальні: комп'ютери/ноутбуки; сервер; офісне приміщення. Фінансові: заробітні плати працівникам; апаратне забезпечення; оренда приміщення.
Головні цілі та завдання проекту	Огляд підходів для статистичного аналізу даних, методів прогнозування та оцінювання ризиків. Створення програмного забезпечення для цієї системи.

## Продовження таблиці 4.1

Очікувані результати	Система автоматизованого моделювання ринкових процесів та оцінювання ризиків стане помічником у багатьох сферах діяльності. Вона дозволить вести більш ефективну економічну діяльність, а також при керуванні бізнесом. Завдяки якісному оцінюванню фінансового ризику компанії зможуть як застрахувати себе від майбутніх збитків, так і виконувати прибуткові ризикові операції з найменшими втратами.
----------------------	--

## 4.3 Команда стартап- проекту

Обов'язки та ключові ролі початкової команди проекту висвітлені у таблиці 4.2.

Таблиця 4.2 – Команда проекту

Посада	Функціональні обов'язки та досвід роботи	Роль
Спеціаліст з аналізу даних	Вибір архітектури проекту; пошук нових галузей в яких доцільно буде використовувати сервіс; статистичний аналіз великих об'ємів даних та пошук зв'язків та закономірностей за допомогою сучасних методів.	Рішення бізнес задач з використанням передових методів обробки даних.
Розробник	Створення програмного забезпечення проекту та вдосконалення існуючих рішень.	Відповідальність за якість та функціональні можливості продукту.
Тестувальник	Пошук імовірних недоліків та помилок у функціонуванні продукту; пропозиції щодо вдосконалення.	Відповідальність за якість та функціональні можливості продукту.

## Продовження таблиці 4.2

Продуктовий менеджер проекту	Бренд-менеджмент; перемовини з замовниками системи; моніторинг та аналіз даного сектору послуг.	Пошук клієнтів, нових ніш на ринку, просування бренду.
------------------------------	---	--

## 4.4 Розробка бізнес-моделі проекту

Бізнес-модель проекту продемонстровано в таблиці 4.3.

Таблиця 4.3 – Бізнес-модель проекту.

Назва	Зміст
Споживчі сегменти	Дослідницькі центри; банки; університети; страхування; енергетика та інші компанії, специфіка роботи яких пов'язана із прогнозуванням та ризиком.

Продовження таблиці 4.3

Ціннісні пропозиції	<p>Ведення більш ефективної економічної діяльності, при керуванні бізнесом.</p> <p>Завдяки якісному оцінюванні фінансового ризику компанії зможуть як застрахувати себе від майбутніх збитків, так і виконувати прибуткові ризикові операції з найменшими втратами.</p> <p>Вирішення складних задач та виявлення основних можливостей для майбутнього розвитку. Продукт, який не потребує особливих налаштувань перед першим використанням.</p>
Канали збуту	<p>Прямі канали:</p> <ul style="list-style-type: none"> <li>— сайт-візитка;</li> <li>— конференції аналітиків/бізнес аналітиків;</li> <li>— презентації для цільового сегменту</li> <li>— розроблених рішень;</li> <li>— презентації в вузах.</li> </ul> <p>Партнерські канали: служба роботи з клієнтами.</p>

Продовження таблиці 4.3

Взаємовідносини з клієнтами	Служба підтримки. Супроводження продукту протягом певного терміну. Особливий підхід для кожного клієнта. Формат договорів на довгосрокову підписку.
Потоки надходження доходу	Розробка спеціалізованого продукту (за умови надання клієнтом бази даних). Покупка ліцензії на певний термін (універсальна web-версія продукту), довгосрокова підписка для використання необхідної частини розробленого рішення; разові виплати за проведення аналізу.
Ключові ресурси	Технічні ресурси, інтелектуальні ресурси, трудові ресурси.
Ключові види діяльності	Аналіз та прогнозування; візуалізація; клієнтська аналітика; керування даними; прийняття рішень; керування ризиками та їх оцінювання.

Продовження таблиці 4.3

Ключові партнери	<p>Компанії, що спеціалізуються на аудиті і консалтингу - актуальна інформація щодо стану економіки, показників, що впливають на прибутки компаній; досвід, що мають компанії в комерційній, технологічній та ін. сферах, а також в проектуванні та поширенні бізнес-рішень.</p> <p>Центри даних - постачання даних для навчання моделей, перевірки їх прогнозної якості, обмін сучасними підходами та моделями.</p> <p>Учбові центри - обмін досвідом, молодими спеціалістами.</p>
Структура витрат	<p>Витрати на зарплатню; оренда серверів; оренда приміщень; відрядження, конференції та внутрішні ресурси; підтримка програмних рішень; просування продукту та пошук нових клієнтів.</p>

#### 4.5 Аналіз ринкових можливостей запуску стартап-проекту

Визначення ринкових можливостей, які можна використати під час ринкового впровадження проекту, та ринкових загроз, які можуть перешкодити реалізації проекту, дозволяє спланувати напрями розвитку проекту із урахуванням стану ринкового середовища, потреб потенційних клієнтів та пропозицій проектів-конкурентів.

Спочатку проводиться аналіз попиту: наявність попиту, обсяг, динаміка розвитку ринку (табл. 4.4).

Таблиця 4.4 – Попередня характеристика потенційного ринку стартап-проекту

Показники стану ринку	Характеристика
Кількість головних гравців, од	3
Динаміка ринку	Зростає
Наявність обмежень для входу	Вихід на ринок прямого конкурента, або суттєве зниження вартості послуг непрямих конкурентів (консалтингових компаній).
Специфічні вимоги до стандартизації та сертифікації	Немає
Середня норма рентабельності в галузі (або по ринку), %	300

Рентабельність — поняття, що характеризує економічну ефективність виробництва, за якої за рахунок грошової виручки від реалізації продукції (робіт, послуг) повністю відшкодовує витрати на її



виробництво й одержується прибуток як головне джерело розширеного відтворення [24].

Суть одного із найважливіших методів оцінки економічної ефективності інвестицій полягає у розрахунку їх середньої рентабельності за формулою [25]:

$$R = \frac{P}{1 + n} * 100, \quad (3.1)$$

де  $P$  - прибуток за час експлуатації проекту;

$n$  - час експлуатації проекту.

Інвестувати грошові засоби доцільно тоді, коли від цього можна отримати більший прибуток, ніж від їх зберігання у банку. Порівнюючи середньорічну рентабельність інвестицій зі ставкою банківського відсотка, можна дійти висновку, що вигідніше [28].

Середня норма рентабельності в галузі (або по ринку) порівнюється із банківським відсотком на вкладення. За умови, що останній є вищим, можливо, має сенс вкласти кошти в інший проект.

За результатами аналізу таблиці робиться висновок щодо того, чи є ринок привабливим для входження за попереднім оцінюванням.

Надалі визначаються потенційні групи клієнтів, їх характеристики, та формується орієнтовний перелік вимог до товару для кожної групи (табл. 4.5).

Таблиця 4.5 – Характеристика потенційних клієнтів стартап-проекту

Потреба, що формує ринок	Цільова аудиторія	Відмінності у поведінці різних потенційних цільових груп	Вимоги споживачів до товару
Необхідне програмне забезпечення для прогнозування показників та оцінювання ринкових ризиків.	Потенційними цільовими групами є дослідницькі центри, університети та компанії, специфіка роботи яких пов'язана із аналізом, ризикменеджментом та прогнозуванням (державний сектор, охорона здоров'я, рітейл, банки).	Відмінність у підходах та застосованих моделей у прогнозуванні та ризикменеджменті, відмінність у сферах діяльності клієнтів.	Рішення має бути швидким, ефективним, мати зрозумілий юзер-френдлі інтерфейс.

Після визначення потенційних груп клієнтів проводиться аналіз ринкового середовища: складаються таблиці факторів, що сприяють ринковому впровадженню проекту, та факторів, що йому перешкоджають (табл. 4.6-4.7). Фактори в таблиці подаються в порядку зменшення значущості.

Таблиця 4.6 – Фактори загроз

Фактор	Зміст загрози	Можлива реакція компанії
Конкуренція	Вихід на ринок великої компанії.	Розширити рекламну кампанію до виходу нового гравця на ринок. Передбачити додаткові переваги власного ПЗ для того, щоб повідомити про них саме після виходу міжнародної компанії на ринок. Обрати нову цільову аудиторію і зосередитися на ній або наявність виходу з ринку
Зміна потреб користувача	Дане ПЗ не задовольняє всі потреби користувачів. Користувачам необхідне ПЗ з іншим функціоналом та підходами до аналізу.	Передбачити можливість додавання нового функціоналу до існуючого рішення. Передбачити можливості зміни напрямку продукту взагалі, так звану точку Pivot.

Продовження таблиці 4.6

Зростання попиту	Занадто швидкі темпи зростання попиту на даного типу ПЗ.	Розширення штату, пошук нових методів для задоволення як найбільшої кількості клієнтів; передбачення гнучкості рішення.
------------------	--	---

Таблиця 4.7 – Фактори можливостей

Фактор	Зміст можливості	Можлива реакція компанії
Конкуренція	Відсутність аналогічного продукту для вітчизняного користувача.	Локалізація та адаптація сервісу для локальних груп. Адаптація до вітчизняних особливостей.
Розробка нових методів прогнозування та оцінювання ризиків.	Розробка нових статистичних методів, що будуть швидші та ефективніші для конкретних цілей.	Покращити ПЗ додаванням нового функціоналу, розширення існуючих можливостей.

Продовження таблиці 4.7

Зростання попиту.	Можливість залучити більшу кількість клієнтів.	Розширення інфраструктури, гнучкість рішення, розширення існуючих можливостей, підходів та моделей.
-------------------	--	---

Надалі проводиться аналіз пропозиції: визначаються загальні риси конкуренції на ринку (табл. 4.8).

Таблиця 4.8 – Ступеневий аналіз конкуренції на ринку

Особливості конкурентного середовища	В чому проявляється дана характеристика	Вплив на діяльність підприємства
Тип конкуренції: монополістична конкуренція	Існує декілька фірм-конкурентів	Підтримка якості продукту та постійні нововведення
За рівнем конкурентної боротьби: міжнародний	Фірми-конкуренти з інших країн	Створити основу ПЗ таким чином, щоб можна було легко переробити дане ПЗ для використання у інших країнах та галузях

Продовження таблиці 4.8

За галузевою ознакою: міжгалузева	Продукт може використовуватись для різних галузей	Постійне вдосконалення продукту, що не має привязки до сфери, гнучкість рішення
За видами товарів: товарно-видова	Види товарів є однаковими: ПЗ для статистики та бізнес- аналізу	Створити ПЗ, враховуючи недоліки конкурентів
За характером конкурентних переваг: нецінова	Вдосконалення технології створення ПЗ, щоб собівартість була нижчою	Використання менш дорогих технологій для розробки, ніж використовують конкуренти
За інтенсивністю: марочна	Велике значення для клієнтів має бренд компаній конкурентів, які вже завоювали місце на ринку	Рекламні стратегії, підтримка компаній партнерів

Таблиця 4.9 – Аналіз конкуренції в галузі за М. Портером

Складові аналізу	Прямі конкуренти в галузі	Потенційні конкуренти	Постачальники	Клієнти	Товари-замінники
	SAS SPSS Eviews	Наявність вже існуючих рішень	—	Контроль якості продукту	Наявність більш широкого функціоналу, зручнішого інтерфейсу та авторитет
Висновки	Доволі інтенсивна конкурентна боротьба з вже закріпившимися на ринку гравцями	Є можливості входу в ринок, але є потенційні конкуренти. Строки виходу на ринок – 6 місяців	—	Клієнти диктують усі умови роботи на ринку	Необхідно випускати ПЗ не гірше, ніж у конкурентів та розширяти функціонал

За результатами аналізу таблиці робиться висновок щодо принципової можливості роботи на ринку з огляду на конкурентну ситуацію. Також робиться висновок щодо характеристик (сильних сторін), які повинен мати проект, щоб бути конкурентоспроможним на ринку. Другий висновок враховується при формулюванні переліку факторів конкурентоспроможності (табл. 4.10).

Таблиця 4.10 – Обґрунтування факторів конкурентоспроможності

Фактор конкурентоспроможності	Обґрунтування
Ціна	Більш доступна ціна збільшує кількість потенційних клієнтів
Функціонал	Методи і моделі, що підходять для багатьох сфер діяльності
Виконання ПЗ у кросплатформеному вигляді	Можливість використання ПЗ на будь-якій платформі
Мобільність	Розробка веб-версій з базовим функціоналом

За визначеними факторами конкурентоспроможності (табл. 4.10) проводиться аналіз сильних та слабких сторін стартап-проекту (табл. 4.11).

Таблиця 4.11 – Порівняльний аналіз сильних та слабких сторін проекту

Фактор конкурентоспроможності	Бали 1-20	Рейтинг товарів конкурентів у порівнянні з даним						
		-3	-2	-1	0	+1	+2	+3
Ціна	10		+					
Функціонал	15				+			
Виконання ПЗ у кросплатформеному вигляді	20					+		
Мобільність	10			+				



Фінальним етапом ринкового аналізу можливостей впровадження проекту є складання SWOT-аналізу (матриці аналізу сильних (Strength) та слабких (Weak) сторін, загроз (Troubles) та можливостей (Opportunities) (табл.4.12) на основі виділених ринкових загроз та можливостей, та сильних і слабких сторін (табл. 4.11).

Таблиця 4.12 – SWOT-аналіз стартап-проекту

Сильні сторони: ціна, функціонал, мобільність	Слабкі сторони: кросплатформеність, недостатня кількість спеціалістів підкованих у найсучасніших методах
Можливості: конкуренція, розробка нових методів прогнозування, зростання попиту	Загрози: конкуренція, зміна потреб користувачів, занадто швидке зростання попиту

На основі SWOT-аналізу розробляються альтернативи ринкової поведінки (перелік заходів) для виведення стартап-проекту на ринок та орієнтовний оптимальний час їх ринкової реалізації з огляду на потенційні проекти конкурентів, що можуть бути виведені на ринок.

Визначені альтернативи аналізуються з точки зору строків та ймовірності отримання ресурсів (табл. 4.13).

Таблиця 4.13 – Альтернативи ринкового впровадження стартап-проекту.

Альтернативи ринкової поведінки	Ймовірність отримання ресурсів	Строки реалізації
Створення повноцінного веб-сервісу	65%	11 місяців
Створення ПЗ	80%	12 місяців
Перехід на безкоштовне розповсюдження	50%	5 місяців

З означених альтернатив обирається та, для якої: а) отримання ресурсів є більш простим та ймовірним; б) строки реалізації – більш стислими.

#### 4.6 Розроблення ринкової стратегії проекту

Розроблення ринкової стратегії першим кроком передбачає визначення стратегії охоплення ринку: опис цільових груп потенційних споживачів (табл. 4.14).

Таблиця 4.14 – Вибір цільових груп потенційних споживачів

Опис профілю цільової групи потенційних клієнтів	Готовність споживачі в сприйняти продукт	Орієнтовний попит в межах цільової групи	Інтенсивність конкуренції в сегменті	Простота входу в сегмент
Компанії (українські та міжнародні), специфіка роботи яких пов'язана із аналізом даних, прогнозуванням та ризикменеджментом	Висока	Високий	Сильна	Складно
Державні підприємства, специфіка роботи яких пов'язана із аналізом даних, прогнозуванням та ризикменеджментом.	Помірна	Високий	Сильна	Складно
Університети та дослідницькі центри, специфіка роботи яких пов'язана із аналізом даних, прогнозуванням та ризикменеджментом.	Помірна	Помірний	Помірна	Середня складність
Як цільові групи обрано: 1 та 3.				

За результатами аналізу потенційних груп споживачів (сегментів) автори ідеї обирають цільові групи, для яких вони пропонуватимуть свій товар, та визначають стратегію охоплення ринку:

— якщо компанія зосереджується на одному сегменті – вона обирає стратегію концентрованого маркетингу;

— якщо працює із кількома сегментами, розробляючи для них окремо програми ринкового впливу – вона використовує стратегію диференційованого маркетингу;

— якщо компанія працює із всім ринком, пропонуючи стандартизовану програму (включно із характеристиками товару/послуги) – вона використовує масовий маркетинг.

Для роботи в обраних сегментах ринку необхідно сформувати базову стратегію розвитку (табл. 4.15).

Таблиця 4.15 – Визначення базової стратегії розвитку

Обрана альтернатива розвитку проекту	Стратегія охоплення ринку	Ключові конкурентоспроможні позиції відповідно до обраної альтернативи	Базова стратегія розвитку
Створення програмного забезпечення для аналізу даних, прогнозування та ризикменеджменту.	Ринкове позиціонування (Позиція компанії чи продукту показує чим він унікальний, чим відрізняється від конкурентів, чим корисний споживачу.)	Простота інтерфейсу, універсальність та ефективність продукту.	Стратегія диференціації.

Наступним кроком є вибір стратегії конкурентної поведінки (табл. 4.16).

Таблиця 4.16 – Визначення базової стратегії конкурентної поведінки

Чи є проект «першопрохідцем» на ринку?	Чи буде компанія шукати нових споживачів, або забирати існуючих у конкурентів?	Чи буде компанія копіювати основні характеристики товару конкурента, і які?	Стратегія конкурентної поведінки
Ні	Як нових, так і вже існуючих.	Так. Проте розробляти більш гнучкі та універсальні методи, розширювати функціонал новими методами та підходами.	Стратегія заняття конкурентної ніші.

На основі вимог споживачів з обраних сегментів до постачальника (стартап-компанії) та до продукту (див. табл. 4.6), а також в залежності від обраної базової стратегії розвитку (табл. 4.15) та стратегії конкурентної поведінки (табл. 4.16) розробляється стратегія позиціонування (табл. 4.17), що полягає у формуванні ринкової позиції (комплексу асоціацій), за яким споживачі мають ідентифікувати торгівельну марку/проект.

Таблиця 4.17 – Визначення стратегії позиціонування

Вимоги до товару цільової аудиторії	Базова стратегія розвитку	Ключові конкурентоспроможні позиції власного стартап-проекту	Вибір асоціацій, які мають сформувати комплексну позицію власного проекту (три ключових)
Ефективність, зручність інтерфейсу, швидкість роботи.	Диференціації	Простота користувацького інтерфейсу, що дозволяє пришвидшити та спростити роботу, швидкість роботи, що дозволяє підвищити швидкість експериментів, ефективність результатів, що дозволяє розробити якісний прогноз та оцінку. гнучкість та мобільність	Стабільність роботи, якість роботи, швидкість роботи, зручність роботи мобільність, гнучкість.

Результатом виконання підрозділу має стати узгоджена система рішень щодо ринкової поведінки стартап-компанії, яка визначатиме напрями роботи стартап-компанії на ринку.

#### 4.7 Розробка маркетингової програми стартап-проекту

Першим кроком є формування маркетингової концепції товару, який отримає споживач.

Таблиця 4.18 – Визначення ключових переваг концепції потенційного товару

Потреба	Вигода, яку пропонує товар	Ключові переваги перед конкурентами
Якість прогнозу оцінки	Якісний прогноз будь-яких показників, визначення ризикових компонент	Відсутність самотійно будувати модель та обирати модель прогнозу. Комплексний підхід до оцінювання ризику.

Продовження таблиці 4.18

Спрощення інтерфейсу користувача	Простота роботи з ПЗ.	Користувачам не потрібно замислюватись над тим, як саме побудувати прогноз. Незалежно, від даних (категоріальних, числових) ПЗ виконає аналіз, запропонує методи та підходи прогнозування та оцінювання ризиків.
Мобільність	Можливість використовувати базову веб-версію.	Користувачі не обов'язково мають бути прив'язані до конкретного місця, а зможуть виконати необхідний базовий аналіз у будь-який час та з будь-якого девайсу.

Надалі розробляється трирівнева маркетингова модель товару: уточнюється ідея продукту та/або послуги, його фізичні складові, особливості процесу його надання (табл. 4.19).



Таблиця 4.19 – Опис трьох рівнів товару

Рівні товару	Сутність та складові		
Товар за задумом	Зручність, швидкість та мобільність отримання практичного результату щодо прогнозування процесів та оцінювання ризиків.		
Товар у реальному виконанні	Властивості/характеристики	М/Нм	Вр/Ех/Тл/Е/Ор
	Якість. Швидкість обробки. Мобільність. Ціна.		
	Якість: функціональне тестування (наявності багів), тестування перформансу, стабільності та глобалізації.		
	Пакування: відсутнє.		
Товар із підкріпленням	1-місячна пробна безкоштовна версія та безкоштовне встановлення на етапі бетатестування.		
	Постійна підтримка для користувачів.		
За рахунок чого потенційний товар буде захищено від копіювання: патент, система захисту.			

Після формування маркетингової моделі товару слід особливо відмітити – чим саме проект буде захищено від копіювання. Захист може бути організовано за рахунок захисту ідеї товару (захист інтелектуальної власності), або ноу-хау, чи комплексне поєднання властивостей і характеристик, закладене на другому та третьому рівнях товару.

Наступним кроком є визначення цінових меж, якими необхідно керуватись при встановленні ціни на потенційний товар (остаточне визначення ціни відбувається під час фінансово-економічного аналізу проекту), яке передбачає аналіз ціни на товари-аналоги або товари

субституту, а також аналіз рівня доходів цільової групи споживачів (табл. 4.20). Аналіз проводиться експертним методом.

Таблиця 4.20 – Визначення меж встановлення ціни

Рівень цін на товари-замінники	Рівень цін на товари-аналоги	Рівень доходів цільової групи споживачів	Верхня та нижня межі встановлення ціни на товар/послугу
50000 грн	1000000 грн	250000 грн	30000 – 40000 грн

Таблиця 4.21 – Формування системи збуту

Специфіка закупівельної поведінки цільових клієнтів	Функції збуту, які має виконувати постачальник товару	Глибина каналу збуту	Оптимальна система збуту
Купують ПЗ та роблять щорічні внески для подовження ліцензії (аналог довгосрокові підписки)	Продаж	0 (напрямку)	Прямий канал збуту

Визначення ідеї та теми рекламного звернення зумовлює всі наступні етапи планування рекламної діяльності. Тому цей етап є основним для

професіоналів — творчих працівників рекламних агенцій чи рекламних підрозділів підприємств.

Таблиця 4.22 – Концепція маркетингових комунікацій.

Специфіка поведінки цільових клієнтів	Канали комунікацій, якими користуються цільові клієнти	Ключові позиції, обрані для позиціонування	Завдання рекламного повідомлення	Концепція рекламного звернення
Купівля ПЗ через Інтернет, робота з ПЗ на комп'ютерах з різними ОС.	Електронна пошта, Інтернет, мобільний зв'язок	Швидкодія, простота використання, ефективність, мобільність.	Показати переваги ПЗ, у тому числі і перед конкурентами	Демо-ролик. Використання сайт-візитки.

#### 4.8 Висновки до розділу

Отже, можливість ринкової комерціалізації проекту присутня, оскільки на даного виду продукти наявний попит, динаміка ринку збільшується, рентабельність роботи на ринку є також значно високою. Існують перспективи впровадження з огляду на потенційні групи клієнтів, бар'єри входження, стан конкуренції, конкурентоспроможність проекту.

Цільовими групами вибрано компанії (українські та міжнародні), специфіка роботи яких пов'язана із аналізом даних, прогнозуванням та ризик-менеджментом. Як альтернативу (варіант) впровадження для ринкової реалізації проекту доцільно обрати розроблення програмного забезпечення. Подальша імплементація проекту є доцільною.

## ВИСНОВОК

В ході проходження практики було досліджено математичні методи і моделі, що дозволяють моделювати складні системи і процеси, оскільки фінансові процеси і є складними системами. Було досліджено величезну кількість джерел на дану тематику.

В даній магістерській дисертації було розглянуто питання оцінки ризику акцій фінансового ринку. Для цього було розібрано методологію оцінювання ризиків. А також методів для прогнозування нестационарних часових рядів. Було оглянуто і досліджено велику кількість джерел на дану тематику. Також було досліджено методи фільтрації даних від шуму і проаналізовано для чого це необхідно. Було розглянуто два методи фільтрації – це фільтр Калмана і адаптивний фільтр Калмана. Було розглянуто методи моделювання гомоскедастичних і гетероскедастичних систем. А також способи оцінки систем на гетероскедастичність.

Звичайно було розглянуто особливості ринкових і зокрема фондових ризиків і методів їх оцінки. А саме методика VaR і використання волатильності для оцінки ризиків.

Також, в ході написання роботи було проведено аналіз даних і пошук оптимальних методів для побудови системи підтримки прийняття рішень для оцінювання ризику акцій.

За результатами досліджень було обрано модель, що дає найкращі результати. І на основі даної моделі було побудовано систему підтримки прийняття рішень. Система була створена на основі емпіричної вибірки даних по цінам на акції компанії Google. Реалізовано систему на мові програмування Python, оскільки дана мова містить велику кількість бібліотек для роботи з часовими рядами і вказаними вище методами. Зокрема зручна робота з нейронними мережами. В кінці було виконано

оцінювання ризику акцій і представлено значення показника VaR для 95% і 99% ймовірностей.

Дану роботу можна і далі продовжувати. До неї можна додати ще методології для оцінки волатильності: нечітку логіку, метод подібних траєкторій, тощо. Також можна узагальнити на теорію портфелів і оцінювати ризики для портфеля акцій.

## ПЕРЕЛІК ПОСИЛАНЬ

1. Turban, E. Decision support and expert systems: management support systems. Englewood Cliffs, N.J.: Prentice Hall, 1995. 976 p.
2. Система підтримки рішень. URL: [https://uk.wikipedia.org/wiki/%D0%A1%D0%B8%D1%81%D1%82%D0%B5%D0%BC%D0%B0\\_%D0%BF%D1%96%D0%B4%D1%82%D1%80%D0%B8%D0%BC%D0%BA%D0%B8\\_%D1%80%D1%96%D1%88%D0%B5%D0%BD%D1%8C](https://uk.wikipedia.org/wiki/%D0%A1%D0%B8%D1%81%D1%82%D0%B5%D0%BC%D0%B0_%D0%BF%D1%96%D0%B4%D1%82%D1%80%D0%B8%D0%BC%D0%BA%D0%B8_%D1%80%D1%96%D1%88%D0%B5%D0%BD%D1%8C) (дата звернення: 03.10.2020)
3. Національний банк України. Ринковий ризик. URL: [https://old.bank.gov.ua/control/uk/publish/article?art\\_id=123616&cat\\_id=123320](https://old.bank.gov.ua/control/uk/publish/article?art_id=123616&cat_id=123320) (дата звернення: 20.10.2020)
4. Power D.J. A Brief History of Decision Support Systems, 2003. URL: <http://dssresources.com/history/dsshhistory.html>
5. Небава М.І. Теорія макроекономіки: Навч. посіб. Київ: Слово, 2005. 536 с.
6. Згуровский М. З., Зайченко Ю. П. Основы вычислительного интеллекта : монография. Київ: Наукова думка, 2013. 406 с.
7. Зайченко Ю.П. Исследование операций: Учеб. Пособие. Київ: Слово, 2003. 688 с.
8. Калина А. В. Современный экономический анализ и прогнозирование (макро- и микроуровень): Учебно-методическое пособие. Київ: МАУП, 1997. 272 с.
9. Бідюк П. І. Курс лекцій по аналізу часових рядів. Київ: НТУУ «КПІ», 2009. 450 с.
10. Ларичев О.И. Теория и методы принятия решения. Москва: Логос, 2000. 296 с.
11. Нейман Дж. Теория игр и экономическое поведение / пер. С англ. Моргенштерн О. Москва: Наука, 1970. 707 с.

12. Little J.D.C. Models and Managers: The Concept of a Decision Calculus. *Management Science*. 1970. Vol. 16. P. B466-B485.
13. Бідюк П.І. Методи прогнозування. Луганськ: Альма Матер, 2008. С. 605.
14. Дрейпер Н. Прикладной регрессионный анализ. Москва: Финансы и статистика, 1986. 366 с.
15. Герасимов Б. М. Человеко-машинные системы принятия решений с элементами искусственного интеллекта. Київ: Наукова Думка, 1993. 184 с.
16. Бидюк П. И. Временные ряды: моделирование и прогнозирование. Киев: ЕКМО, 2003. 141 с.
17. Бокс Дж., Анализ временных рядов. Прогноз и управление. Москва: Мир, 1974. 402 с.
18. Бідюк П.І. Принципи прогнозування часових рядів. *Наукові вісті НТУУ „КПІ”*. 2005. № 5. С. 14-25.
19. Gustafsson. Particle filter theory and practice with positioning applications. *Aerospace and Electronic Systems Magazine*. 2010. Vol. 25. P. 53–82.
20. Brown R.G. Smoothing forecasting and prediction if discrete time series. New York: Courier Corporation, 1963. 468 p.
21. Бідюк П.І. Проектування комп'ютерних інформаційних систем підтримки прийняття рішень: Навчальний посібник. Київ: ННК «ІПСА» НТУУ «КПІ», 2010. 340 с.
22. Кириченко О. Банківський менеджмент : навчальний посібник для вищих навчальних закладів. Київ : ОСНОВИ, 1999. 671 с
23. Штучна нейронна мережа. URL: [https://uk.wikipedia.org/wiki/%D0%A8%D1%82%D1%83%D1%87%D0%BD%D0%B0\\_%D0%BD%D0%B5%D0%B9%D1%80%D0%BE%D0%BD%D0%BD%D0%B0\\_%D0%BC%D0%B5%D1%80%D0%B5%D0%B6%D0%B0](https://uk.wikipedia.org/wiki/%D0%A8%D1%82%D1%83%D1%87%D0%BD%D0%B0_%D0%BD%D0%B5%D0%B9%D1%80%D0%BE%D0%BD%D0%BD%D0%B0_%D0%BC%D0%B5%D1%80%D0%B5%D0%B6%D0%B0)



24. Рентабельність виробництва і методика визначення її показників. URL: <http://buklib.net/books/29473/>

25. Прогнозування ефективності інвестиційного проекту. URL: [http://pidruchniki.com/1566072162240/turizm/prognozuvannya\\_efektivnosi\\_investitsiynogo\\_proektu](http://pidruchniki.com/1566072162240/turizm/prognozuvannya_efektivnosi_investitsiynogo_proektu)

26. Розробка та перевірка концепції товару. URL: <http://westudents.com.ua/glavy/35910-3-rozrobka-ta-perevrkakontsepts-tovaru.html>

27. Концепції маркетингової діяльності. URL: <http://referat-ok.com.ua/marketing/marketingova-diyalnist-2>

## ДОДАТОК А. ЛІСТИНГ ПРОГРАМИ

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
import statsmodels.api as sm
import arch
from statsmodels.graphics.tsaplots import plot_acf
from numpy import exp,sqrt
from scipy.stats import norm
import scipy.stats as scs
import statsmodels.tsa.api as smt
import os
import sys
import pandas as pd
#import pandas_datareader.data as pdr
import numpy as np
import statsmodels.formula.api as smf
import statsmodels.tsa.api as smt
import statsmodels.api as sm
import scipy.stats as scs
from arch import arch_model
import matplotlib.pyplot as plt
import matplotlib as mpl
%matplotlib inline

data = pd.read_csv('Google.csv', delimiter=';')
```

```
data.index = list(data['Date'])
```

```
forecast_horizon = 30
```

```
df = data[['Date', 'Close']]
```

```
df['daily_returns'] = np.log(df.Close / df.Close.shift(1)) * 100      # Daily log
returns
```

```
df['monthly_std'] = df['daily_returns'].rolling(21).std()  # Standard deviation
across trading month
```

```
df['annual_vol'] = df['monthly_std'] * np.sqrt(252)      # Annualize monthly
standard deviation
```

```
df = df.dropna().reset_index(drop=True)
```

```
train = df[:-forecast_horizon]
```

```
train.index = list(train['Date'])
```

```
test = df[-forecast_horizon:]
```

```
test.index = list(test['Date'])
```

```
describe_stat = scs.describe(train.Close)
```

```
# Kalman filter
```

```
from pykalman import KalmanFilter
```

```
data.Close.fillna(0, inplace=True)
```

```
kf = KalmanFilter(initial_state_mean=0)
```

```
state_means,_ = kf.filter(data.Close.values)
```

```
state_means = state_means.flatten()
```

```
indexDate=data.index
```

```
new_data=pd.Series(state_means,index=indexDate)
```

```
plt.plot(data.Close.values[:300])
plt.plot(new_data.values[:300])
plt.show()
```

```
df = pd.DataFrame(new_data, columns = ['Close'])
df['Date'] = new_data.index
df['daily_returns'] = np.log(df.Close / df.Close.shift(1)) * 100      # Daily log
returns
df['monthly_std'] = df['daily_returns'].rolling(21).std() # Standard deviation
across trading month
df['annual_vol'] = df['monthly_std'] * np.sqrt(252)      # Annualize monthly
standard deviation
df = df.dropna().reset_index(drop=True)
```

```
train = df[0:len(df)-forecast_horizon]
train.index = list(train['Date'])
test = df[-forecast_horizon:-1]
test.index = list(test['Date'])
```

```
describe_stat = scs.describe(test.Close)
```

```
def tsplot(y, lags=None, figsize=(10, 8), style='bmh'):
    if not isinstance(y, pd.Series):
        y = pd.Series(y)
    with plt.style.context(style):
        fig = plt.figure(figsize=figsize)
        layout = (3, 2)
        ts_ax = plt.subplot2grid(layout, (0, 0), colspan=2)
```

```

acf_ax = plt.subplot2grid(layout, (1, 0))
pacf_ax = plt.subplot2grid(layout, (1, 1))
qq_ax = plt.subplot2grid(layout, (2, 0))
pp_ax = plt.subplot2grid(layout, (2, 1))
y.plot(ax=ts_ax)
ts_ax.set_title('Time Series Analysis Plots')
smt.graphics.plot_acf(y, lags=lags, ax=acf_ax, alpha=0.5)
smt.graphics.plot_pacf(y, lags=lags, ax=pacf_ax, alpha=0.5)
#sm.qqplot(y, line='s', ax=qq_ax)
#qq_ax.set_title('QQ Plot')
#kde_pdf = scs.gaussian_kde(y)
#x= np.linspace(min(y), max(y), len(y))
#bins=np.arange(min(y), max(y),0.5)
#pp_ax.plot(x,scs.norm.pdf(x,scs.tmean(y),scs.tstd(y)),label="parametric
distribution",color="red")
#pp_ax.plot(x,kde_pdf(x),'r',label="KDE estimation",color="blue")
#pp_ax.hist(y,bins=bins,density = True,color="gray",alpha=.6, label='data
histogram')
#pp_ax.set_title("Probability density function")
pp_ax.legend()
plt.tight_layout()
plt.show()
return

_=tsplot(train.Close, lags=12)

am=arch_model(train.Close)
res = am.fit(dispen='off')
print(res.summary())

```

```
import statsmodels.stats.diagnostic as smd
```

```
smd.het_arch(res.resid)
```

```
_=tsplot(res.resid**2, lags=12)
```

```
#GARCH
```

```
am=arch_model(train.Close)
```

```
res = am.fit(dispatch='off')
```

```
print(res.summary())
```

```
#heterosked test
```

```
import statsmodels.stats.diagnostic as smd
```

```
smd.het_arch(res.resid)
```

```
scs.describe(res.resid)
```

```
resid_all = res.resid/res.conditional_volatility
```

```
scs.describe(resid_all)
```

```
def mae(y_true, y_pred):
```

```
    y_true, y_pred = np.array(y_true), np.array(y_pred)
```

```
    return np.mean(np.abs(y_true - y_pred))
```

```
def mape(y_true, y_pred):
```

```
    y_true, y_pred = np.array(y_true), np.array(y_pred)
```

```

return np.mean(np.abs((y_true - y_pred) / y_true))

def theil(y_true, y_pred):
    y_true, y_pred = np.array(y_true), np.array(y_pred)
    y_true_sq = np.sqrt(np.sum(y_true**2)/y_true.size)
    y_pred_sq = np.sqrt(np.sum(y_pred**2)/y_pred.size)
    y_sq = np.sqrt(np.mean((y_true - y_pred)**2))
    return y_sq/(y_true_sq+y_pred_sq)

def mdl_conclusion(res):
    res_ser = pd.Series()
    res_ser = pd.Series(res.params).append([res.pvalues, pd.Series(res.model,
index=['MDL']),pd.Series([res.nobs, res.loglikelihood, res.aic, res.bic],
index=['Num', 'LL', 'AIC', 'BIC'])])
    return res_ser

from arch.univariate import ConstantMean, GARCH, Normal, StudentsT

am = ConstantMean(train.Close)
am.volatility = GARCH(p=12, q=100, power=2)
#am.distribution = StudentsT()
am.distribution = Normal()
res = am.fit(dispen='off')
print(res.summary())

end_loc = 1500
res_test = am.fix(res.params.values, last_obs=end_loc)
pred_test = res_test.forecast(horizon=1, align = 'origin').variance
print(pred_test.shift(1)[end_loc:])

```

```

sample_var = res.conditional_volatility[end_loc:]
forecasted_var = pred_test.shift(1)[end_loc:]
#forecasted_var = pd.DataFrame(forecasts).T['h.1']
#df = pd.concat([sample_var,forecasted_var],1)
#df.columns = ['actual', 'forecasted']
plt.plot(sample_var, label = 'actual')
plt.plot(forecasted_var, label = 'forecasted')
plt.legend()
plt.show()

print('mae:', mae(sample_var, forecasted_var))
print('mape:', mape(sample_var, forecasted_var))
print('theil:', theil(sample_var, forecasted_var))

#egarch = EGARCH(p=1, q=1)
#egarch = EGARCH(p=1, o=1, q=1)
#exponential arch process: earch = EGARCH(p=5)
from arch.univariate import EGARCH
am = ConstantMean(train.Close)
am.volatility = EGARCH(p=3, o=1, q=12)
am.distribution = Normal()
#am.distribution = StudentsT()
res = am.fit(dispen='off')
print(res.summary())

end_loc = 1500
res_test = am.fix(res.params.values, last_obs=end_loc)
pred_test = res_test.forecast(horizon=1, align = 'origin').variance

```



```
print(pred_test.shift(1)[end_loc:])
```

```
sample_var = res.conditional_volatility[end_loc:]
```

```
forecasted_var = pred_test.shift(1)[end_loc:]
```

```
#forecasted_var = pd.DataFrame(forecasts).T['h.1']
```

```
#df = pd.concat([sample_var, forecasted_var], 1)
```

```
#df.columns = ['actual', 'forecasted']
```

```
plt.plot(sample_var, label = 'actual')
```

```
plt.plot(forecasted_var, label = 'forecasted')
```

```
plt.legend()
```

```
plt.show()
```

```
print('mae:', mae(sample_var, forecasted_var))
```

```
print('mape:', mape(sample_var, forecasted_var))
```

```
print('theil:', theil(sample_var, forecasted_var))
```

```
am = ConstantMean(train.Close)
```

```
am.volatility = GARCH(p=3, o=1, q=12, power=1.0)
```

```
am.distribution = StudentsT()
```

```
#am.distribution = Normal()
```

```
res = am.fit(dis='off')
```

```
print(res.summary())
```

```
end_loc = 1500
```

```
res_test = am.fix(res.params.values, last_obs=end_loc)
```

```
pred_test = res_test.forecast(horizon=1, align = 'origin').variance
```

```
print(pred_test.shift(1)[end_loc:])
```

```

sample_var = res.conditional_volatility[end_loc:]
forecasted_var = pred_test.shift(1)[end_loc:]
#forecasted_var = pd.DataFrame(forecasts).T['h.1']
#df = pd.concat([sample_var,forecasted_var],1)
#df.columns = ['actual', 'forecasted']
plt.plot(sample_var, label = 'actual')
plt.plot(forecasted_var, label = 'forecasted')
plt.legend()
plt.show()

```

```

print('mae:', mae(sample_var, forecasted_var))
print('mape:', mape(sample_var, forecasted_var))
print('theil:', theil(sample_var, forecasted_var))

```

####VAR

```

def VaR_gaus(p, volatility, mean):

```

```

    v_a_r=pd.Series()
    coef = scs.norm.ppf(p)
    v_a_r=mean + coef*volatility
    return v_a_r

```

```

def ES_gaus(p, volatility, mean):

```

```

    es=pd.Series()
    coef = -scs.norm.pdf(scs.norm.ppf(1-p))/(p)
    es=mean + coef*volatility
    return es

```

```
def VaR_t(p, volatility, mean, nu):
```

```
    v_a_r=pd.Series()
```

```
    coef = np.sqrt((nu-2)/nu)*scs.t.ppf(p, df = nu)
```

```
    v_a_r = mean + coef * volatility
```

```
    return v_a_r
```

```
def ES_t(p, volatility, mean, nu):
```

```
    es=pd.Series()
```

```
    coef = np.sqrt((nu-2)/nu)*(-scs.t.pdf(scs.t.ppf(1-p,df = nu), df =nu)/p)*((nu +
(scs.t.ppf(p, df = nu))**2)/(nu-1))
```

```
    es=mean + coef * volatility
```

```
    return es
```

```
#quantile for 95% (also need for 99%)
```

```
p=0.05
```

```
-scs.norm.pdf(scs.norm.ppf(1-p))/(p)
```

```
#alpha = 1-p = 0.95
```

```
nu=1000
```

```
np.sqrt((nu-2)/nu)*(-scs.t.pdf(scs.t.ppf(1-p,df = nu), df =nu)/p)*((nu +
(scs.t.ppf(p, df = nu))**2)/(nu-1))
```

```
import tensorflow as tf
```

```
from tensorflow.keras.models import Sequential
```

```
from tensorflow.keras.layers import Dense, Dropout, LSTM
```

```
import numpy
```

```
import matplotlib.pyplot as plt
```

```
import pandas
```

```

import math

from sklearn.preprocessing import MinMaxScaler
from sklearn.metrics import mean_squared_error

# fix random seed for reproducibility
numpy.random.seed(7)

# MinMaxNormalization

dataset = forecasted_var['h.1'].apply(lambda x: (x - min(forecasted_var['h.1'])) /
(max(forecasted_var['h.1']) - min(forecasted_var['h.1'])))

target = sample_var.apply(lambda x: (x - min(sample_var)) / (max(sample_var)
- min(sample_var)))

train_size = int(len(dataset) * 0.7)
test_size = len(dataset) - train_size

train = dataset[0:train_size]
test = dataset[train_size:len(dataset)]

target_train = target[0:train_size]
target_test = target[train_size:len(target)]

len(train), len(test), len(target_train), len(target_test)

# convert an array of values into a dataset matrix
def create_dataset(dataset, look_back=1):
    dataX, dataY = [], []
    for i in range(len(dataset)-look_back-1):

```

```

    a = dataset[i:(i+look_back)]
    dataX.append(a)
    dataY.append(dataset[i + look_back])
return numpy.array(dataX), numpy.array(dataY)

look_back = 1
trainX, trainY = create_dataset(list(target_train), look_back)
testX, testY = create_dataset(list(target_test), look_back)

# reshape input to be [samples, time steps, features]
trainX = numpy.reshape(trainX, (trainX.shape[0], 1, trainX.shape[1]))
testX = numpy.reshape(testX, (testX.shape[0], 1, testX.shape[1]))

# create and fit the LSTM network
model = Sequential()
model.add(LSTM(4, input_shape=(1, look_back)))
model.add(Dense(1))
model.compile(loss='mean_squared_error', optimizer='adam')
model.fit(trainX, trainY, epochs=100, batch_size=1, verbose=2)

# make predictions
trainPredict = model.predict(trainX)
testPredict = model.predict(testX)

trainPredict = trainPredict.reshape(1, len(trainPredict))[0]
testPredict = testPredict.reshape(1, len(testPredict))[0]

pred_train = []

```

```
pred_test = []
```

```
y_train = []
```

```
y_test = []
```

```
trainScore = math.sqrt(mean_squared_error(trainY, trainPredict))
```

```
print('Train Score: %.2f RMSE' % (trainScore))
```

```
testScore = math.sqrt(mean_squared_error(testY, testPredict))
```

```
print('Test Score: %.2f RMSE' % (testScore))
```

```
trainScore = math.sqrt(mean_squared_error(y_train, pred_train))
```

```
print('Train Score: %.2f RMSE' % (trainScore))
```

```
testScore = math.sqrt(mean_squared_error(y_test, pred_test))
```

```
print('Test Score: %.2f RMSE' % (testScore))
```

```
print('mae:', mae(y_train, pred_train))
```

```
print('mae_test:', mae(y_test, pred_test))
```

```
print('mape:', mape(y_train, pred_train))
```

```
print('mape_test:', mape(y_test, pred_test))
```

```
print('theil:', theil(y_train, pred_train))
```

```
print('theil_test:', theil(y_test, pred_test))
```

```
p = 0.01
```

```
nu = 1000
```

```
from scipy.stats import norm
```

```
var = []
```

```
var_t = []
```

```
for i in range(len(pred_test)):
    mu = np.mean(data)
    vol = pred_test[i]
    #var.append(norm.ppf(confidence_level=p, mu, vol))
    var.append(VaR_gaus(p, vol, mu))
    var_t.append(VaR_t(p, vol, mu, nu))
```